

Many thanks to:



National Science Foundation
WHERE DISCOVERIES BEGIN



NYU Institute of Human Development
and Social Change (IHDSC)

and Cybele Raver, Director of NYU IHDSC

*data coding, analysis, archiving, and
sharing for open collaboration:*

From **open**SHAPA to Open Data Sharing



September 15-16, 2011 (Thu. & Fri.)

Stafford Place II
4121 Wilson Boulevard, 5th Floor
Arlington, VA, 22230

Data Coding, Analysis, Archiving, and Sharing for Open Collaboration: From OpenSHAPA to Open Data Sharing

Researchers now have access to richer and more detailed behavioral data than ever before. For example, when studying how children learn to walk, researchers can collect eye-tracking data from miniature head-mounted cameras recording the infant's eye movements and field of view, making it possible to see exactly where the child looks while navigating through the environment. Simultaneously, researchers can collect high-speed motion-tracking data detailing the trajectories of the child's limb movements and video data about the child's path relative to caregivers and obstacles, interactions with people, objects, and surfaces, and affective responses while walking, falling, and interacting. Despite the widespread availability of video and other recording technologies, behavioral researchers typically settle for analyzing only one variable in one stream of data, rather than seeking relations among multiple variables across multiple data streams. Powerful data analysis tools and sophisticated data management practices are needed to integrate different kinds of data and relate them to each other—tools and practices that few researchers have. In addition, researchers usually work in isolation, seldom sharing data that might illuminate others' research. Without richer analyses and data sharing, theoretical progress in developmental psychology and other fields of behavioral science is hampered.

The purpose of this workshop is to delve into the conceptual, technical, and management issues that, when resolved, will allow researchers to perform richer analyses across large, shared, data sets. The workshop will focus in part on the future development of an emerging open-source software tool, OpenSHAPA, and will explore how OpenSHAPA might be extended to encompass new data exploration and visualization tools and promote data management and data sharing. Twenty-two researchers will participate in the workshop, representing the fields of cognitive, perceptual, social, language, and motor development, human-computer interaction, visual analytics, computer science, eResearch, cognitive science, and human factors. Collectively, the invited researchers have experience with different aspects of the problem of exploring rich behavioral data, such as performing massive data visualization, innovative data analyses, integrating multiple data streams, performing custodianship of shared data sets, and creating eResearch communities and data management tools.

The outcomes from the workshop will help to improve the quality of behavioral science. First, findings from the workshop will have an immediate impact on further development of the OpenSHAPA tool, where development is shared across a burgeoning community of users. Possible directions are changes to the architecture to prepare for expansion of data management and data sharing capabilities, building links to existing software, creating libraries of scripts for users to manage data in standardized ways, creating web-based user guides and best practices, expanding user forums, and providing efficient technical support. Research community members can freely adopt OpenSHAPA, expand their current use of it, or build bridges between it and other open source tools, and will bring new users into the community of current users and developers. Second, the richer data analysis that results should support richer theoretical insights. Better data management practices will support more reliable and replicable research, and will better preserve data for future use within and across laboratories. A community of open data sharing practices will lead to greater transparency and efficiency in research and teaching by allowing researchers to inspect each other's data sets and analyses, thereby reducing puzzling failures to replicate, generating new hypotheses, and exposing students to original footage of tasks and findings.



Eric Zhang

Email: eez206@nyu.edu Mobile: 301-693-3636

Workshop Contact Person

Day 1 (Thursday, September 15)

8:00-8:30	<i>Breakfast</i>
8:30-10:00	Introduction Karen Adolph, Penelope Sanderson, Clinton Freeman, Jesse Lingeman
10:00-10:30	<i>Morning break</i>
10:30-12:00	Data Coding, Management & Sharing Linda Smith, John Stamper, Brian MacWhinney
12:00-1:30	<i>Lunch</i>
1:30-3:00	Data Mining, Visualization, Analysis, Plug-Ins Daniel Messinger, Chen Yu, Katy Borner
3:00-3:30	<i>Afternoon break</i>
3:30-5:30	Data Sharing: Professional & Technical Issues Marc Bornstein, Bennett Bertenthal, Micah Altman, Pamela Davis-Kean
7:00	<i>Workshop dinner at Willow</i> 4301 N. Fairfax Drive, Arlington, VA 22203

Day 2 (Friday, September 16)

8:00-8:30	<i>Breakfast</i>
8:30-10:00	Managing Multiple Data Streams Martha Alibali, Mike Goldstein, Cole Galloway
10:00-10:30	<i>Morning Break</i>
10:30-12:00	Data Annotation, Exploration, Visualization Robert Hoffman, Wayne Gray, William Wong
12:00-12:15	Identification of 3-5 key issues for discussion after lunch
12:15-1:30	<i>Lunch</i>
1:30-3:00	Problem-oriented group discussions
3:00-3:30	<i>Afternoon break</i>
3:30-4:00	Short summary presentations from group discussions
4:00-5:00	Summary Richard Aslin, Francis Quek, Jeff Lockman
7:00	<i>Workshop dinner at Rock Bottom</i> 4238 Wilson Boulevard, Ste. 1256, Arlington, VA 22203

Organizers



Karen E. Adolph

Email: karen.adolph@nyu.edu

KAREN E. ADOLPH is Professor in the Department of Psychology and the Center for Neuroscience at New York University. She received her B.A. from Sarah Lawrence College in 1986 and her Ph.D. from Emory University in 1993, and completed a postdoctoral fellowship at the Albert Einstein College of Medicine in 1994. Her first faculty position was at Carnegie Mellon University (1994-1997). Dr. Adolph has received a James McKeen Cattell Sabbatical Award, the Robert L. Fantz Memorial Award from the American Psychological Foundation, the Boyd McCandless Award from the American Psychological Association, the Young Investigator Award from the International Society for Infant Studies, FIRST and MERIT awards from the National Institute of Child Health and Human Development, two Distinguished Teaching awards from the Psychology Department at NYU, and the Golden Dozen and Distinguished Teaching Awards from NYU. She is a Fellow of the American Psychological Association and the Association for Psychological Science. Her research has been continually supported by grants from the National Institutes of Health since 1996. She chaired the NIH study section on Motor Function and Speech Rehabilitation, is on the Advisory Board of the McDonnell Foundation, and is on the editorial boards of *Developmental Psychobiology* and *Ecological Psychology*. Dr. Adolph is author of the monograph *Learning in the Development of Infant Locomotion*, multiple handbook chapters on motor development, and 92 other publications.

Primary research interests

Adolph's work is inspired by a developmental systems approach and her research interests include learning and development in the context of motor skill acquisition. She is especially interested in effects of body growth, exploratory activity, environmental and social supports, and culture on perceptual-motor learning and development. Her research—in collaboration with more than 170 high school students and undergraduates and 54 doctoral students, postdocs, and research staff—examines sitting, crawling, walking, brachiation, postural control, reaching, and object manipulation in infant humans and monkeys, children, young adults, and elderly adults using state-of-the-art video recording, motion tracking, eye tracking, and gait carpets. With Penelope Sanderson, she has spearheaded the development of the OpenSHAPA tool. She helped to innovate the first head-mounted eye-tracking system for use in mobile infants and children.

Challenges for Open Data Sharing

- A tool that is flexible enough to suit the needs of individual researchers and labs yet powerful enough to support a collaborative community of behavioral researchers
- Data management issues such as linking coding criteria with spreadsheets, recording coders/dates/discrepancy resolution, inter-rater reliability, managing scripts for manipulating data and implementing plug-ins
- Linking multiple streams of behavioral data (video, physiological, etc.)
- Visualizing and exploring data in terms of sequence, co-occurrence, and other patterns
- Tackling professional and technical issues in data sharing across labs (IRB and parental consent, assigning credit to PIs and funding agencies, limiting use)



Penelope Sanderson

Email: psanderson@nicta.com.au

PENELOPE SANDERSON leads the Cognitive Engineering Research Group at The University of Queensland, where she is Professor of Cognitive Engineering and Human Factors in the Schools of Psychology, Information Technology and Electrical Engineering, and Medicine. She is also Project Leader of a major cognitive and organisational systems engineering initiative within NICTA, Australia's leading ICT research and development organisation. Sanderson's principal research contributions relate to the design and evaluation of human-system integration for complex sociotechnical systems. She received her BA(Hons) from University of Western Australia in 1979 and her MA and PhD from University of Toronto in 1981 and 1985. She was on the faculty of University of Illinois at Urbana-Champaign from 1985 to 1996 before returning to Australia in 1997. Sanderson is a Fellow of the Academy of the Social Sciences in Australia. In 2004 she was awarded the Distinguished International Colleague Award from the US-based Human Factors and Ergonomics Society (HFES) and in 2009 she was a Royal Academy of Engineering Distinguished Visiting Fellow in the UK. Sanderson has twice received the HFES Jerome H. Ely Award (1990 and 2005) for the best paper in HFES's peer-reviewed journal, *Human Factors*, and she has won eight best conference paper awards with her students. She is currently on the editorial boards of *Cognition, Technology, and Work*; *Journal of Experimental Psychology: Applied*; *Human Factors*; and *Journal of Cognitive Engineering and Decision Making*.

Primary research interests

Sanderson investigates human-system integration in complex work environments, such as critical care units in hospitals, air traffic control centres, and power system control centres. Her focus is how best to support the human role in such environments, through better management and display of information. Examples of recent work performed with colleagues and students include the design of auditory displays and head-mounted displays to help clinicians monitor the status of anesthetized patients; exploration of a "distributed prospective memory" approach to understanding the impact of interruptions and reminders in busy work environments; and the development of human and automated reasoning methods to predict the impact of new eHealth technologies on work patterns and information flow in healthcare contexts. Sanderson's theoretical approach is influenced by cognitive work analysis, distributed cognition, and European work psychology, and her investigations usually take place in the field or in high-end simulators. Key forms of data collected are audiovisual recordings of activity and verbalization, process status logs, and eye-tracking data. Sanderson directed the development of the SHAPA software for verbal protocol analysis and the MacSHAPA software for video analysis, and with Carolanne Fisher she developed the Exploratory Sequential Data Analysis framework. Together with Karen Adolph she is overseeing development of OpenSHAPA.

Challenges for Open Data Sharing

- Rapid assembly, integration, and synchronization of multiple sources of data in multiple media.
- Unobtrusive but powerful support for a broad range of hermeneutic analyses, from loose interpretation to tightly-controlled coding schemes, with smooth transitions over that range
- Maintaining links across all levels and all stages of analysis, for easy revisiting of original data such as audiovisual recordings and of steps undertaken during analysis
- Ability to rapidly adjust exploratory sequential data analysis software and/or adapt and integrate other software to meet ever-evolving research needs
- Managing the software and scripting contributions of a community of researcher/developers



Clinton Freeman

Email: clinton.freeman@nicta.com.au

CLINTON FREEMAN is a Senior Research Engineer at National ICT Australia (NICTA). He received his BCST(Hons) from the University of Sydney in 2003, before working as a software engineer for the Arts eResearch unit at the University of Sydney. During his time with the Arts eResearch unit he developed software to support the digital curation and dissemination of large archaeological datasets of images, sounds and maps.

From 2005, Clinton spent several years building software in industry. While, working for Gemcom Software International he worked on 3D visualisations of open pit mines and novel algorithms for dealing with distortion in mapping aerial photography on virtual terrains.

Since 2008, Clinton has been involved in the development of OpenSHAPA. Putting in place tools and processes to create an environment from which the first Alpha and Beta versions of OpenSHAPA emerged. He designed and developed the foundations of important elements such as the scripting, plugin and spreadsheet architectures that others are now using to build amazing things.

Primary Research interests

Freeman is interested in understanding the messy support tasks that hinder and slow academic research. He is a practitioner of iterative and agile software development approaches to quickly prototype and evolve robust software. Clinton has also built:

- The values viewer – A desktop tool for visualizing the impact of technological changes on various stakeholders of a system.
- UserMetrix [<http://usermetrix.com>] -- a telemetry tool for software developers that measures the impact on an application's user experience – to help him understand and build better tools for supporting research.

Challenges for Open Data Sharing

- Research is groundbreaking and new. Tools for supporting research need to be flexible and easily customizable, so they can be tailored to different research situations.
- Labs and researchers are drowning under an ever-increasing volume of digital data; we need better ways to support the high level management across projects and experiments.
- Research data from different digital sources need to be easily combined, visualized and observed.
- It is still difficult to collaborate and share digital research data across geographically dispersed labs and institutes.



Jesse M. Lingeman

Email: jmlingeman@nyu.edu

JESSE LINGEMAN is a Research Scientist in Prof. Karen Adolph's Action Lab and a Computer Science Masters Student at New York University. He received his BA from Western Michigan University in 2008. Jesse's focus in computer science is on machine learning, specifically learning latent structures within data. In the Action Lab, he has created software to record and analyze motion capture data, gait carpet data, and distance sensor data. He is designing the interface for a robot chair for studies of passive locomotion and rotation. His developmental research focuses on infant walking, brachiation in children and adults, and reaching in infant monkeys.

Jesse's Masters Thesis is about inferring how genes regulate each other from time-series data. He has created a software package to infer and analyze gene regulatory networks using an ensemble of other algorithms, and an accompanying textbook for beginning bioinformaticians.

Jesse has also been heavily involved in developing the OpenSHAPA Ruby Scripting Interface and OpenSHAPA plugins. The Ruby Scripting Interface allows users to easily manipulate their data in OpenSHAPA using simple Ruby commands. The plugins will allow for new visualizations alongside the OpenSHAPA spreadsheet, including a timeline view and 3D motion capture data.

Email: jmlingeman@nyu.edu

Primary research interests

Jesse Lingeman is interested in understanding the hidden structures of data, and figuring out new ways to learn from data. This may take the form of developing new algorithms to look for structures, figuring out how to make a robot interact with its environment, or coming up with a new way to visualize data.

Participants



Martha W. Alibali

Email: mwalibali@wisc.edu

MARTHA W. ALIBALI is Professor in the Departments of Psychology and Educational Psychology at the University of Wisconsin-Madison. She received her B.A. from University of Chicago in 1986 and her Ph.D. from University of Chicago in 1994. Her first faculty position was at Carnegie Mellon University (1994-1999). Alibali has received the Robert L. Fantz Memorial Award from the American Psychological Foundation (2004), a Vilas Associate award from the University of Wisconsin (2006), and the Hamel Family College of Letters and Sciences Faculty Fellow award from the University of Wisconsin (2009). She has also received the Chancellor's Distinguished Teaching Award (2004), and she is a Fellow of the American Psychological Society. Her research has been supported by grants from the National Institutes of Health, the National Science Foundation, and the United States Department of Education Institute for Education Sciences. She served as a member of the NIH study section on Cognition and Perception for four years, and she currently serves as a member of the IES Basic Processes review panel. She serves on the editorial boards of *Developmental Psychology*, *Cognition and Instruction*, and the *Journal of Cognition and Development*. Alibali is co-author (with Robert Siegler) of the cognitive development textbook *Children's Thinking* as well as over 90 other publications.

Primary research interests

Alibali's research investigates processes of knowledge change in cognitive development and mathematics learning. She investigates both conceptual and procedural aspects of students' developing knowledge, and how these forms of knowledge are used in reasoning and problem solving. A central focus of the work is on the role of manual gestures in learning and communication, particularly in educational settings. Her prior work has shown that students' gestures reveal important information about their mathematical reasoning processes--information that is not revealed in their speech. Current projects explore how teachers use gestures in classroom settings, how teachers' gestures influence students' learning, and how perception, action, and gesture are involved in mathematical thinking and learning.

Challenges for Open Data Sharing

- Data exploration and visualization, particularly for classroom and tutorial interactions
- IRB and consent issues, especially as they apply to classroom data
- Integrating multiple data streams, e.g., activities of the teacher and multiple small groups of students in a classroom
- Integrating quantitative and qualitative measures in data analysis
- Integrating across different video software tools, e.g., ELAN and Transana



Micah Altman

Email: micah_altman@alumni.brown.edu Web: <http://redistricting.info>

MICAH ALTMAN is Senior Research Scientist in the Institute for Quantitative Social Science in the Faculty of Arts and Sciences at Harvard University and Archival Director of the Henry A. Murray Research Archive.

Dr. Altman has authored over 50 publications and 5 open-source software packages that span informatics, statistics, computer science, political science, and other social-science disciplines. His extensively-reviewed book, *Numerical Issues in Statistical Computing for the Social Scientist*, corrects common computational errors made across the range of social sciences. Dr. Altman has won the 2009 Best Political Science Research Software Award, honorable mention for the Library Technology Excellence Award, the 2005 Best Political Science Software Award and the 1999 Best Political Science Research Website Award, given by the American Political Science Association (CMS & ITP), the 1999 Outstanding Dissertation Award, given by the Western Political Science Association, for best dissertation in the Western region; and the 1998 Weaver Award, for best paper in representation and electoral systems, given by the American Political Science Association.

Dr. Altman is on the editorial board of *The Journal of Information Technology and Politics* and has served on the board of *The International Encyclopedia of Political Science* and *The Journal of E-Government*; as a guest editor for *The Social Science Computer Review*, and *Journal of Statistical Software*; as the head of Information Technology and Politics section of the Midwestern Political Science Association conference, and as the president of the Information Technology and Politics organized section of the American Political Science Association.

Primary research interests

Dr. Altman conducts research in social science informatics, social science research methodology, and American politics, focusing on the intersection of information, technology, and politics; and on the dissemination, preservation, and reliability of scientific knowledge. His current research interests include computationally reliable and efficient statistical methods; the collection, sharing, citation and preservation of research data; the creation and analysis of networks of scientific knowledge; information privacy; and computational methods of redistricting.

Challenges for Open Data Sharing

- In data sharing, issues of legal ownership, scholarly attribution, research provenance, direct access, and long-term persistence are often entangled. A major challenge is to separate these concerns and developing independent but interoperating mechanisms for storage, citation, tracking, discovery that are sustainable in the scholarly communication ecosystem.
- In the social sciences and humanities, data may contain private information about individuals. None of the licensing frameworks, and few technical frameworks support effective sharing of such data, especially across institutional and jurisdictional boundaries.
- Systems for coordinating the collection and analysis data such as Kepler and VisTrails are becoming increasingly more sophisticated, but lack any connection or support for data sharing and publication. Integrating the sharing of data and publication of results into the same workflow environments can both reduce the effort and increase quality the quality of both research and traditional publications.



Richard N. Aslin

Email: aslin@cvs.rochester.edu

RICHARD N. ASLIN is the William R. Kenan Jr. Professor in the Department of Brain and Cognitive Sciences at the University of Rochester. He is also Director of the Rochester Center for Brain Imaging. He received his B.A. from Michigan State University in 1971 and his Ph.D. from the University of Minnesota in 1975. His first faculty position was at Indiana University (1975-1984). Dr. Aslin has received the Boyd McCandless Award and the Early Career Award in developmental psychology from the American Psychological Association, a John Simon Guggenheim Fellowship, an honorary doctorate from Uppsala University, and the Goergen Award for Contributions to Undergraduate Learning at the University of Rochester. He is the current president of the International Society for Infant Studies and is a Fellow of the American Academy of Arts & Sciences. His research has been supported by grants from the National Institutes of Health, the National Science Foundation, and the McDonnell Foundation. He is the past editor of *Infancy* and currently is on the editorial boards of *Cognition*, *Cognitive Science*, and *Language Learning & Development*. He has over 100 publications in refereed journals and many book chapters.

Primary research interests

Aslin's research is focused on the rapid learning of sequences of events in the auditory domain and spatial configurations of visual elements in complex scenes. This statistical learning mechanism, which is both highly susceptible to environmental inputs and constrained by innate biases, enables adults, children, and infants to group sounds based solely on the distributional information (conditional probabilities) of elements contained in the sound stream or in visual scenes. He and his collaborators have also used a serial reaction time task to study the on-line processing of sequential statistics and fMRI to study the neural correlates of statistical learning. A feature of his lab since 1978 has been the use of eye-tracking to study visual attention, and this work continues using Bayesian models of familiarity and novelty preferences. He has also spearheaded the use of a new brain imaging technique, called Near-Infrared Spectroscopy, that can assess the neural correlates of language and visual processing, including statistical learning, in young infants. Finally, his lab is using a head-mounted camera and eye-tracker to study how infants view natural scenes, particularly during language learning.

Challenges for Open Data Sharing

- Merging multiple data streams of time-synchronized eye-tracking data
- Extracting correlations among patterns of looking, talking, and gesturing
- Visualizing the predictors and consequences of dyadic interactions



Bennett I. Bertenthal

Email: bbertent@indiana.edu

BENNETT I. BERTENTHAL is the James H. Rudy Professor of Psychological and Brain Sciences at Indiana University. He served as dean of the College of Arts and Sciences from 2007 - 2010. Prior to this appointment, he was Professor of Psychology and Computational Neuroscience at the University of Chicago, and a Senior Fellow of the Computation Institute. He served as the Assistant Director of the Social, Behavioral, and Economic Sciences (SBE) Directorate of the National Science Foundation (NSF) from 1996 - 1999. Dr. Bertenthal began his career at the University of Virginia in 1979 and remained there for 20 years. He received his B.A. from Brandeis University in 1971 and his Ph.D. from the University of Denver in 1978, and was a postdoctoral fellow at the Brain Research Institute at UCLA from 1978-1979. Dr. Bertenthal is the author of over 120 publications on the development of perception and action, developmental cognitive neuroscience, visual processing of motion information, perception and production of biological motions, multimodal communication, and science infrastructure and policy. During his career he has received numerous federal and private foundation research, training, and infrastructure grants, and was the founding Director of the NSF funded Social Informatics Data Grid (SIDGrid). He has lectured extensively on the social and behavioral sciences and science policy in the U.S. and abroad, and has served on various national advisory committees including the National Science and Technology Council subcommittee concerned with basic science and fundamental research on children. He is a fellow of the American Association for the Advancement of Science, the American Psychological Society, and the American Psychological Association.

Primary research interests

My research focuses on the origins, development, and basic processing mechanisms involved in perception, action, and cognition. Special emphasis is currently directed toward the coupling between the observation and execution of actions and how this perception-behavior link contributes to early development. More generally, the goal of this research is to understand the dynamic interplay between the development of the brain, the body, and the mind in the social and physical world. This research agenda requires familiarity and experience with a multidisciplinary set of research methods and techniques such as video-based coding of behavior, machine-coding of behavior, data visualization techniques, eye tracking, biomechanics, signal processing, psychophysics, electrophysiology, and chronometric approaches. These methods result in huge data sets that require linear and nonlinear modeling. The research is supported by a large network of collaborators whose expertise complements the mission of the lab. In order to accomplish these goals, it has been necessary to develop centers and infrastructure projects, such as the Social Informatics Data Grid (SIDGrid).

Challenges for Open Data Sharing

- Develop software tools for collecting and analyzing large corpuses of audio, video, and physiological data
- Develop a distributed data repository for storing, retrieving and searching streaming data files with the support of an ensemble of web and grid tools
- Provide a Science Gateway to the Teragrid for analyzing huge or computationally intensive data sets
- Establish standards for annotating, coding, and analyzing shared data
- Engage research community via education, training, and outreach



Katy Börner

Email: katy@indiana.edu

KATY BÖRNER is the Victor H. Yngve Professor of Information Science at the School of Library and Information Science, Adjunct Professor at the School of Informatics and Computing, Adjunct Professor at the Department of Statistics in the College of Arts and Sciences, Core Faculty of Cognitive Science, Research Affiliate of the Biocomplexity Institute, Fellow of the Center for Research on Learning and Technology, Member of the Advanced Visualization Laboratory, and Founding Director of the Cyberinfrastructure for Network Science Center (<http://cns.iu.edu>) at Indiana University. She is a curator of the Places & Spaces: Mapping Science exhibit (<http://scimaps.org>). Her research focuses on the development of data analysis and visualization techniques for information access, understanding, and management. She is particularly interested in the study of the structure and evolution of scientific disciplines; the analysis and visualization of online activity; and the development of cyberinfrastructures for large scale scientific collaboration and computation. She is the co-editor of the Springer book on '*Visual Interfaces to Digital Libraries*' and of a special issue of *PNAS* on 'Mapping Knowledge Domains' (2004). Her new book *Atlas of Science: Visualizing What We Know* by MIT Press was published in 2010. She holds a MS in Electrical Engineering from the University of Technology in Leipzig, 1991 and a Ph.D. in Computer Science from the University of Kaiserslautern, 1997.



Marc H. Bornstein

Email: bornstem@mail.nih.gov

MARC H. BORNSTEIN is Senior Investigator and Head of Child and Family Research at the Eunice Kennedy Shriver National Institute of Child Health and Human Development. He has held faculty positions at Princeton University and New York University and visiting professorships and fellowships around the globe (Munich, London, Paris, Tokyo, Cameroon, Seoul, Trento, Santiago) to support his cross-cultural research. He holds a B.A. from Columbia College, Ph.D. from Yale University, and an honorary doctorate from the University of Padua. Bornstein was a J. S. Guggenheim Foundation Fellow, and he received a Research Career Development Award from NICHD. He also received the C. S. Ford Cross-Cultural Research Award from the Human Relations Area Files, the Boyd McCandless Young Scientist Award and the G. Stanley Hall Award from the American Psychological Association, a United States PHS Superior Service Award from the NIH, two Japan Society for the Promotion of Science Fellowships, three Awards for Excellence from the American Mensa Education & Research Foundation, the Arnold Gesell Prize from the Theodor Hellbrügge Foundation, MERIT award from the NIH, and the Distinguished International Contributions to Child Development Award from the Society for Research in Child Development. Bornstein is on the Executive Committee of the International Society of Infancy Studies and a past member of the Governing Council of the Society for Research in Child Development. He was named to the Top 20 Authors for Productivity in Developmental Science by the American Educational Research Association.

Bornstein is coauthor of *Development in Infancy* (5 editions), *Development: Infancy through Adolescence*, *Lifespan Development*, and *Perceiving Similarity and Comprehending Metaphor*. He is general editor of *The Crosscurrents in Contemporary Psychology Series* and he is general editor of the *Monographs in Parenting* series. He edited *Maternal Responsiveness: Characteristics and Consequences*, the *Handbook of Parenting* (Vols. I-V, 2 editions), and *The Handbook of Cultural Developmental Science* (Parts 1 & 2), and has coedited *Developmental Science: An Advanced Textbook* (6 editions), *Stability and Continuity in Mental Development*, *Contemporary Constructions of the Child*, *Early Child Development in the French Tradition*, *The Role of Play in the Development of Thought*, *Acculturation and Parent-Child Relationships*, *Immigrant Families in Contemporary Society*, *The Developing Infant Mind*, and *Early Childhood Development and Later Achievement*. He is author of or consultant on several children's books, videos, and puzzles in *The Child's World* and *Baby Explorer* series.

Bornstein administers both Federal and Foundation grants, sits on the editorial boards of several professional journals, is a member of scholarly societies in a variety of disciplines, and consults for governments, foundations, universities, publishers, scientific journals, the media, and UNICEF. Bornstein is Editor Emeritus of *Child Development* and founding Editor of *Parenting: Science and Practice*. He has published in experimental, methodological, comparative, developmental, and cultural science as well as neuroscience, pediatrics, and aesthetics. Visit www.cfr.nichd.nih.gov and www.tandfonline.com/HPAR.



Pamela E. Davis-Kean

Email: pdakean@umich.edu

PAMELA E. DAVIS-KEAN is an Associate Professor in the Department of Psychology at the University of Michigan and a Research Associate Professor at the Institute for Social Research. She received her B.A. from Florida State University in 1987 and her Ph.D. from Vanderbilt University in 1996, and completed a postdoctoral fellowship at the Survey Research Center at the Institute for Social Research where she designed the Child Development Supplement of the Panel Study of Income Dynamics. She is currently the Director and PI of the NSF supported Collaboration on Analysis for Pathways from Childhood to Adulthood (CAPCA) which replicates developmental studies across multiple national, international, and community studies. Her research has been continually supported by grants from the National Institutes of Health and the National Science Foundation since 1996. She has authored handbook chapters on parenting and development methods, and over 30 journal articles chapters on parent's educational attainment and the home environment as well as self-esteem development.

Primary research interests

My research focuses on the influence of parent's educational attainment on the construction of home environments; the importance of looking at race and culture when understanding families and child development; the role that families, schools, and significant figures play (multi-context) in the development of children; the development of self-esteem across the lifespan; and gender and math achievement. I am also interested in the psychometric properties of questionnaires and large-scale data collection of data on families and children.

Challenges for Open Data Sharing

- How to convince users the importance of using shared data resource.
- Dealing with issues related to authorship when using collaborative data.
- How to form research teams for using the data in novel and unique ways.
- Why systems like OpenSHAPA are important for our science to validate research questions and to replicate research findings.



James C. Galloway

Email: jacgallo@udel.edu

JAMES C. GALLOWAY is Professor in the Department of Psychology and the Center for Neuroscience at New York Dr. Galloway is currently an Associate Professor, Dept of Physical Therapy, University of Delaware. His research team includes parents, early educators, therapists, industry partners, physicians as well as clinical and engineering researchers. His lab is currently funded by multiple grants from NIH and NSF.

Primary research interests

Since 2000, his Infant Behavior Lab has focused on better understanding how infants use their head, arm and leg movements to physically explore their world. Most recently, he has focused on interventions that assist young children with special needs in maximizing their exploration and the effect of this increased exploration on cognition, language and socialization. Current projects include assessments of brain-behavior relationships in preterm infants, family-based interventions and reality-based technologies including low and high tech power mobility options for infants and toddlers.

Challenges for Open Shapa

- Merging streams of various levels of data within the research subjects (ex. video, behavioral coding, muscle activity patterns, motion data) and outside the subject (ex. sensors in environment).
- Data mining large multiple stream data sets.
- Real time aspects of data gathering and analysis. Example: such as automatic event or pattern recognition during behaviors between one or more subjects.
- Ability to better work with data from distant sites. Example the ability to code behavioral data located at distant site.



Michael H. Goldstein

Email: mhg26@cornell.edu

MICHAEL H. GOLDSTEIN is Associate Professor in the Department of Psychology at Cornell University. He is a member of the graduate fields of Psychology, Cognitive Science, and Human Development. He received his B.A. from University of Colorado – Denver in 1993 and his Ph.D. from Indiana University in 2001. His first faculty position was at Franklin & Marshall College (2001 – 2004). Dr. Goldstein has received the David Kucharski Young Investigator Award from the International Society for Developmental Psychobiology, the Distinguished Early Career Contribution Award from the International Society on Infant Studies, and the Robert A. and Donna P. Paul Award for Excellence in Advising from Cornell University. He is on the editorial boards of *Developmental Psychobiology* and *Parenting: Science and Practice*. His research is currently supported by grants from NICHD and NSF. Dr. Goldstein's publications have appeared in top-tier journals such as *Proceedings of the National Academy of Sciences* and *Psychological Science*.

Primary research interests

Goldstein studies the development and evolution of vocal learning. He takes a comparative approach, investigating the ontogeny of vocal communication in human infants and songbirds. He focuses on the developmental processes by which knowledge of speech and song is acquired from the social environment. To investigate the processes by which vocal development is constructed from interactions with caregivers, he takes a micro-analytic approach to social learning, observing and manipulating parent-offspring interaction in real-time to study development as it proceeds over multiple timescales from seconds to months. Dr. Goldstein's research, which has involved 96 undergraduates and 8 doctoral students, has found that immature vocalizations (plastic song in birds and babbling in human infants) create opportunities for learning from parental feedback to those sounds, and thus have important functional significance in the development of mature song and speech. He is developing an ecological theory of communicative development that is grounded in basic mechanisms of sensory development and social learning.

Challenges for Open Data Sharing

- Integrating streams of behavioral and physiological data from several individuals in a family over multiple timescales
- Flexibly defining co-occurrence and contingency windows to assess interaction patterns
- Exporting acoustic data to specialized programs (e.g. Praat, Sound Analysis Pro) for analysis and re-importing results back to the main timeline, while maintaining temporal synchrony with other data types
- Establishing a large set of high-quality audio/video records of songbird vocal development in social context that will be shared by multiple researchers, each with their own research questions, in the same spirit as Brian MacWhinney's CHILDES project.



Wayne D. Gray

Email: grayw@rpi.edu

WAYNE D. GRAY is a researcher in the fields of computational cognitive modeling, cognitive neuroscience, interactive behavior, cognitive task analysis, cognitive workload, and human error. Since earning his Ph.D. from UC Berkeley he has worked for government and industry research laboratories, as well as universities. He is currently a Professor of Cognitive Science at Rensselaer Polytechnic Institute. Wayne is a Fellow of the Cognitive Science Society, the Human Factors & Ergonomics Society (HFES), and the American Psychological Association (APA). In 2008, APA awarded him the Franklin V. Taylor Award for Outstanding Contributions in the Field of Applied Experimental & Engineering Psychology. He is a past Chair of the Cognitive Science Society and the founding Chair of the Human Performance Modeling technical group of HFES. At present he is a Consulting Editor for the Psychological Review and the Executive Editor for the Cognitive Science Society's first new journal in 30 years, Topics in Cognitive Science (topiCS). Most recently he has been elected a Fellow by the Alexander von Humboldt Foundation to pursue research at the Max Planck Institute Center for Adaptive Behavior and Cognition (ABC) in Berlin.

Primary research interests

Wayne seeks to understand how goal-directed cognition is shaped by the accommodation of basic interactive routines to the cost structure of the task environment. Interactive routines are composites of cognitive, perceptual, and motor operators that come together at the 1/3 to 3 second timespan to form elementary but meaningful units of interactive behavior. They are posited to form the basic behavioral units of interactive behavior for human-device, human-information, and human-human interactions. The selection of these basic elements of integrated behavior is typically beneath our conscious awareness and deliberate control. Hence, the claim (not uncommon in Cognitive Science) is that non-deliberate forces that dynamically react to our task environment shape a large part of our mental life.

There is a basic and applied component to this research agenda. The Cognitive Science side focuses on the control of interactive behavior, resource allocation, dynamic decision-making, memory, attention, and motor movement. The Cognitive Engineering side can be characterized by the terms visual-analytics, human-computer interaction, cognitive workload, and human error. The two types of research feed into each other and are supported by a core of common techniques and methods including computational cognitive modeling, cognitive task analysis, and detailed collection and analysis of behaviors that take less than 1000 milliseconds to occur (e.g., eeg, keystrokes, mouse movements, and eye gaze) and which occur in near instantaneous response to dynamic system events.

Challenges for Open Data Sharing

- Tools to control the flow of data from the lab, to the journal paper, and beyond
- Tools that will aggregate sampled data to form meaningful units at different levels of analysis; e.g., sampled eye data (240 hz) into eye fixations (≈ 10 to 5 hz) that can be combined with motor data (e.g., mouse clicks) and external events to form interactive routines (≈ 3 to 0.3 hz).
- High bandwidth data collection with well formatted records, easy to reuse documentation, and ability to address new questions after the data is collected
- Automated or semi-automated protocol analyses, which enable theory-based parsing of log files to form runnable cognitive models.
- Visualizing and exploring data in terms of sequence, co-occurrence, and other patterns



Robert R. Hoffman

Email: rhoffman@ihmc.us

Web: www.ihmc.us/groups/rhoffman/

ROBERT R. HOFFMAN, Ph.D. is a Senior Research Scientist at the Institute for Human and Machine Cognition in Pensacola Florida. Hoffman is a recognized world leader in cognitive systems engineering and Human-Centered Computing. He is a Fellow of the Association for Psychological Science and a Fulbright Scholar. His Ph.D. is in experimental psychology from the University of Cincinnati, where he received McMicken Scholar, Psi Chi, and Delta Tau Kappa Honors. Following a Postdoctoral Associateship at the Center for Research on Human Learning at the University of Minnesota, he joined the faculty of the Institute for Advanced Psychological Studies at Adelphi University. He began his career as a psycholinguist, and founded the journal, *Metaphor and Symbol*. His subsequent research leveraged the psycholinguistics background in the study of methods for eliciting the knowledge of domain experts. Hoffman has been recognized internationally in disciplines including psychology, artificial intelligence, remote sensing, and weather forecasting, for his research on human factors, the psychology of expertise, the methodology of cognitive task analysis, human-centering issues intelligent systems technology and the design of macrocognitive work systems. Hoffman is a Co-Editor for the Department on Human-Centered Computing in *IEEE: Intelligent Systems*. He is Editor for the book Series, "Expertise: Research and Applications." He is a co-founder and member of the Board of Advisors for the *Journal of Cognitive Engineering and Decision Making*.

Primary research interests

His current research focuses on methodological issues in the analysis of complex systems, and performance measurement and metrics for macrocognitive work systems. He is spearheading the development of a new branch of statistics called Range Statistics, to address some outstanding measurement problems including the problem of "small numbers of small numbers" and the problem of assessing practical significance. With IHMC colleagues, he is developing a new methodology for task analysis of human-robot team collaboration.

Challenges for Open Data Sharing

- Frame Problems, including the problem of a priori versus ad hoc categories.
- Envisioned World problems, including the need to design for adaptivity.
- Practical problems of representation and coding.
- Human-centering issues including the need to design for kluges and the need to design based on desirements rather than requirements.



Jeff J. Lockman

Email: lockman@tulane.edu

JEFFREY J. LOCKMAN is Professor in the Department of Psychology at Tulane University. He received his B.A. from New York University in 1974 and his Ph.D. from the Institute of Child Development, University of Minnesota in 1980. He is Director of Psychology Graduate Studies and past Chair of the Psychology Department at Tulane. He is currently Editor of *Child Development* and has previously served as an Associate Editor for *Child Development*. He is a Fellow of the Association for Psychological Science. His research has been supported by grants from the National Science Foundation and the National Institutes of Health. Additionally, he served as an inaugural member of the NIH study section on Motor Function and Speech Rehabilitation. His publications center on the development of motor behavior and spatial cognition and was co-editor with Esther Thelen of a special section on *Developmental Biodynamics: Brain, Body, Behavior Connections* published in *Child Development*.

Primary research interests

Lockman's research has focused on the development of manual skill and spatial cognition. In recent work, he has been developing a perception-action approach to the study of tool use. In this approach, he has been examining how objects function as extensions of the hand and how early object manipulation serves as a substrate for the emergence of tool use. His work in this area combines the use of behavioral and kinematic methodologies and instrumentation, involving high-speed motion tracking, eye tracking and force plates. Among the manual skills that he and his students are studying are hammering, object fitting, drawing and chopstick use.



Brian MacWhinney

Email: macw@cmu.edu

BRIAN MACWHINNEY is Professor of Psychology, Computational Linguistics, and Modern Languages at Carnegie Mellon University. He received a Ph.D. in psycholinguistics in 1974 from the University of California at Berkeley. He has spent sabbatical years as a Visiting Distinguished Professor at Hong Kong University, Southern Denmark University, and the Hong Kong Institute of Education. MacWhinney's work with Elizabeth Bates on crosslinguistic studies of language processing led to the formulation of the Competition Model during the 1980s (MacWhinney & Bates, 1989). The Competition Model has been applied to the study of normal adult sentence processing, the development of sentence processing in childhood, language in children with focal lesions, second language acquisition, specific language impairment, and language processing in aphasia. This work has led to neural network simulations for the learning of the German declensional system, English verb morphology, lexical categorization, conceptual prototypes, childhood bilingualism, and first language vocabulary acquisition. More recently, MacWhinney has incorporated the core assumptions of the Competition Model into a Unified Competition Model that links various forms of language learning and processing to emergentist theories, including the theory of embodied cognition (MacWhinney, 1999).

In 1984, he and Catherine Snow co-founded the CHILDES (Child Language Data Exchange System) Project for the computational study of child language transcript data. The CHILDES programs and database have now become an important component of the basic methodology of research in language acquisition. In 2011, he received the first Roger Brown Award for this work in child language acquisition methodology and theory. He is now extending this system to seven additional research areas in the form of the TalkBank Project: AphasiaBank, TBIBank, DementiaBank, FluencyBank, PhonBank, CABank, and SLABank. MacWhinney's recent work includes studies of online learning of second language vocabulary and grammar, neural network modeling of lexical development, fMRI studies of children with focal brain lesions, and ERP studies of between-language competition. He is also exploring the role of grammatical constructions in the marking of perspective shifting, the determination of linguistic forms across contrasting time frames, and the construction of mental models in scientific reasoning.

Challenges for Open Data Sharing

- Full commitment to data-sharing mandated by funding agencies
- Interoperability
- XML Schema compatibility
- Open data access



Daniel Messinger

Email: dmessinger@miami.edu

DANIEL MESSINGER is Professor in the Department of Psychology at the University of Miami with secondary assignments in Pediatrics and Electrical and Computer Engineering. He received a B.A. from Haverford College, a M.A. from the University of Chicago, and a Ph.D. from the University of Utah. He has received awards from the International Society for Infant Studies and the Science Directorate of the American Psychological Association. Dr. Messinger's work—conducted in collaboration with multidisciplinary teams of postdocs, and graduate and undergraduate students—is funded by the NIH, NSF, and Autism Speaks. He is director of the Miami Marino Autism Research Institute and is an Executive Committee Member of the NIH/Autism Speaks Baby Siblings Research Consortium. He was associate Editor of the journal *Emotion* and was on the editorial board of *Infancy*. He is a grant reviewer for Autism Speaks, the NIH, and the NSF. Dr. Messinger is the author of 42 peer-reviewed journal articles and 15 invited chapters and encyclopaedia articles.

Primary research interests

Dr. Messinger is interested in identifying the emergence of basic competencies in ongoing interactions that enable later developmental achievements. His theoretical orientation is dynamic systems theory. He investigates infants' development of fundamental emotional and social competencies, and their impact on a variety of outcomes. This work is conducted from a developmental psychopathology perspective. Dr. Messinger asks how typical developmental processes are disrupted in psychopathology, and then addresses what these disruptions reveal about typical development. He works with typically developing infants, infants at risk for autism spectrum disorders, children with autism, and children with prenatal drug exposure. Dr. Messinger is particularly interested in real-time emotional interaction and its role in development, with emphases on facial expression and gaze shifting in parent-child interaction, and security of attachment. His is also collecting motion capture and video data of infant social and motor development to inform the design of a developing robot. Dr. Messinger is active in developing and disseminating efficient, objective measurement approaches with which to better understand interaction and foster pathways to optimal development. Tools and resources include analytic methods (e.g., windowed cross-correlations, bootstrapping of event sequences), observational methods (e.g., real-time continuous rating and coding software), automated measurement (e.g., computer vision and pattern recognition), and video recordings (see <http://measurement.psy.miami.edu/>).

Challenges for Open Data Sharing

- Building a scientific culture among behavioral scientists, reviewers, IRBs, and participants that values sharing observational phenomena and datasets (e.g., audiovisual recordings and their measurements).
- Increasing the visibility and accessibility of flexible tools for measuring observational behavior (e.g., manual coding and automated measurement).
- Increasing awareness of and infrastructure for the appropriate analysis of observational (primarily audiovisual) data (from ascertaining coder reliability to bootstrapping sequence patterns).



Francis Quek

Email: quek@vt.edu

FRANCIS QUEK is a Professor of Computer Science at Virginia Tech. He also directs the Vision Interfaces and Systems Laboratory. Francis received both his B.S.E. summa cum laude (1984) and M.S.E. (1984) in electrical engineering from the University of Michigan. He completed his Ph.D. C.S.E. at the same university in 1990. Francis is a member of the IEEE and ACM. Professor Quek has previously been on the faculty of the University of Illinois at Chicago and Wright State University, and on the research faculty of the University of Michigan at Ann Arbor. Professor Quek has received the Excellence in Research award at Wright State University. He has authored over 160 research publications in the areas of Computer Vision, Human-Computer Interaction, and Medical Imaging.

Primary research interests

He performs research in embodied interaction, multimodal verbal/non-verbal interaction, physical computing, interactive multimodal meeting analysis, vision-based interaction, multimedia databases, medical imaging, collaboration technology, human computer interaction, computer vision, and computer graphics. He leads several multiple-disciplinary research efforts to understand the communicative realities of multimodal interaction.

Challenges for Open Data Sharing

- Support for 'situated analysis' such that researchers can drill into specific instances of a behavior and view/analyze it within context. This includes contextual time-synchronous viewing.
- Support for multiple synchronized stream (eye-gaze direction, hand motion, EEG etc.), ordinal transcription (manual transcriptions, automatically detected events etc.), and media (audio, video) data.
- Support for 'situated tagging' for phenomena within data such that observations, insights, and questions can be shared.
- Support for synchronous multi-time-scale analysis. This includes data synchronization and synchronous analysis at scales other than video rate (e.g., time-synchrony at millisecond levels to investigate audio/speech phenomena).
- Support for interaction with multiple tools: e.g. csv, Praat, Transcriber, Anvil, MacVisSTA, Matlab, SPSS ...
- Supporting seamless integration across different analysis platforms, individual analysis, and enterprise-level shared analysis and coding
- Support for interactive behavioral model building, search, and situated access to behavior instances
- Support for shared coding, inter-coder reliability, etc.
- Support for 'citable data' (e.g. Dataset xxxyyy, time index: tt, Stream: sss1, sss2, sss3 ... to broaden sharing.)



Linda Smith

Email: smith4@indiana.edu

LINDA SMITH is a Distinguished Professor and the Chancellor's Professor of Psychological and Brain Sciences and Cognitive Science at Indiana University - Bloomington. She received her B.S. degree in 1973 from the University of Wisconsin - Madison and her Ph.D. in psychology from the University of Pennsylvania in 1977. She joined the faculty at Indiana University in 1977 and currently serves as Chair of the Department of Psychological and Brain Sciences. She has published over 200 research articles and is a fellow of the American Academy of Arts and Sciences, the Society of Experimental Psychologists, Cognitive Science and the American Psychological Society. You may find out more about her research and laboratory at www.iub.edu/~cogdev.

Primary research interests

Smith's research is directed to understanding developmental process especially interactions among perception, action and language as children break into language learning. Her current work focuses on the role of sensory-motor processes (eye, head, and hand coordination) in visual attention and on interrelations among visual and manual object exploration, the development of object recognition, and early word learning in late infancy. With the goal of understanding how the coordination of multiple systems in a real time task support and/or create developmental change, she is using multi-modal measures of sensory-motor systems as infants engage with objects and as they learn object names. These measures include head-mounted cameras, multiple video and audio measures, motion tracking, and eye-tracking systems.

Challenges for Open Data Sharing

- Pattern discovery is the key challenge. The advances in real-time multimodal measures yield massive amounts of data at a level of analysis not yet studied. How do discover the regularities at this new level?
- What are the proper kinds of statistics? Psychology has been committed to statistics across individuals, with the subject as the random variable. But within these dense multi-modal data sets, these kinds of analyses seem not right. Information theoretic approaches, time-series analyses, and modeling all might offer better insights, but does the field need statistical advances in these areas with respect to developmental questions, and measures of change in developing systems?
- When does one share data? How long is appropriate for a laboratory to have proprietary rights over a data set? Who gets to use the data? And who decides?
- Do need to develop conventions for data storage and notation so that it can be minded from many different frameworks with many different goals?



John C. Stamper

Email: john@stamper.org

JOHN C. STAMPER is a member of the research faculty at the Human-Computer Interaction Institute at Carnegie Mellon University, and also the Technical Director of the Pittsburgh Science of Learning Center DataShop. John earned his PhD in Information Technology from the University of North Carolina at Charlotte in 2010, an MBA from the University of Cincinnati in 2000, and a BS in Systems Analysis from Miami University in 1994. Prior to returning to academia, John spent over ten years in the software industry where he managed the development and design of complex database systems for large businesses. John is a Microsoft Certified Systems Engineer (MCSE) and a Microsoft Certified Database Administrator (MCDBA). He was the co-chair of the 2010 KDD Cup Competition, titled “Educational Data Mining Challenge,” which was a high profile data mining competition centered on improving assessment of student learning. John was a keynote speaker at the 2011 Educational Data Mining conference where he discussed the growth in educational data and the need to build open repositories to share these data.

Primary research interests

Stamper’s work is centered on using data collected from educational systems to improve student learning. One line of research is the automatic generation of intelligent tutoring capabilities via educational data mining. The Hint Factory is a novel application of this approach and is used to automatically generate contextualized hints and feedback from past student data. Another area of research is the use of data mining to improve cognitive models of student knowledge. These models drive many of the instructional decisions that intelligent tutors make, including how to organize instructional messages, the sequence of topics, and problem selection in a curriculum. Traditionally, cognitive task analysis (CTA) is used to create these models, but CTA is expensive and the models created often do not adequately fit the data. Data mining techniques can suggest improvements to these models which can improve the overall efficiency of student learning leading to a significant savings in time needed for students to learn skills. John’s research is enabled by DataShop, which is an open data repository and set of associated visualization and analysis tools. DataShop has data from thousands of students derived from interactions with on-line course materials and intelligent tutoring systems. The data is fine-grained, with student actions recorded roughly every 20 seconds, and it is longitudinal, spanning semester or yearlong courses. As of August 2011, over 280 datasets are stored including over 60 million student actions which equates to over 175,000 student hours of data. Most student actions are “coded” meaning they are not only graded as correct or incorrect, but are categorized in terms of the hypothesized competencies or knowledge components needed to perform that action. DataShop allows researchers to import data in order to use the provided analysis tools, and to export data from the repository to perform additional analysis. Researchers have analyzed these data to better understand student cognitive and affective states and the results have been used to redesign instruction and demonstrably improve student learning.

Challenges for Open Data Sharing

- Building data models that are flexible for a broad range of researchers.
- Security and Privacy issues related to data sharing
- Managing the risks of secondary analysis (misunderstanding the data)
- Handling the explosive growth of data available



William Wong

Email: w.wong2@btopenworld.com

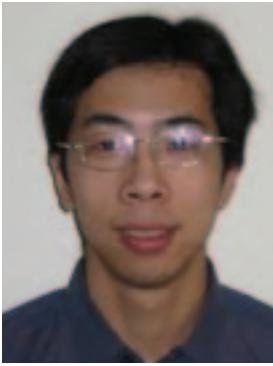
WILLIAM WONG is Professor of Human-Computer Interaction and Head, Interaction Design Centre, at Middlesex University's School of Engineering and Information Sciences, London, UK. Prior to academia, he worked 13 years in the military, first as an infantry platoon commander and then as an air defence controller in the Republic of Singapore Air Force. Later, as Staff Officer for Command Support and Special Projects Officer, he was responsible to CAF for wartime command and control information management. In his final Air Force posting as Head, Systems and Communications Operations Branch, HQ Republic of Singapore Air Force, he set up the unit for the wartime command and control of electronic information assets. In 1992 he joined the Department of Information Science, University of Otago, New Zealand, as Assistant Lecturer and part-time PhD candidate. He later established and directed the Multimedia Systems Research Laboratory and was promoted to Associate Professor in 2003. He completed his PhD at Otago in 1998, entitled, "Information portrayal in dynamic decision making environments" in which he researched how decision strategies can be identified and supported with appropriate information designs, in the naturalistic decision making environment of emergency ambulance command and control in New Zealand and Australia; later in England, Scotland and Wales, and other dynamic environments such as air traffic control, hydro-electricity dispatching, and in learning and evaluation. He is currently investigating the problems of visual analytics in sense-making domains with high information density and variability. Recipient of over £4.5 mil in grants, he currently leads an European Commission FP7 funded project, "CRISIS", that has 12 research and industrial partners across 7 countries, investigating a new interactive simulation environment for training crisis managers at airports; is deputy PI for a 9-partner EPSRC-funded Making Sense visual analytics-oriented project; and is Coordinator for the UK Visual Analytics Consortium (UKVAC), jointly funded by the US Department of Homeland Security, and the UK HM Government's Home Office and Ministry of Defence.

Primary research interests

- Representation design and naturalistic decision making in the context of human-computer interaction and cognitive systems engineering
- Cognitive Task Analysis, using the Critical Decision Method and extending it to speed up and systematize the data analysis process, called Emergent Themes Analysis.
- Visual Analytics: designing the visual representations and interactivity for sense-making in various information environments, simplifying yet communicating multi-dimensional relationships, develop system architectures and conceptual frameworks for technical and operational integration of 'smart' algorithms

Challenges for Open Data Sharing

In May 2009, Wong and Sanderson, through a Royal Academy of Engineering Distinguished Visiting Fellowship, visited the Swedish Defence Research Agency FOI where they were presented with the problem of needing to rapidly analyse very large and mixed collections of performance data for After Action Review. These problems include (Wong and Varga, 2011), the problem of: (a) seeing a large data set and reasoning space through a small keyhole; (b) aggregating and reconciling multiple points of view or predictions; (c) evidence collation and evidential reasoning; (d) provenance and traceability of analytic reasoning; (e) integrating data space, analytic space and hypothesis spaces; (f) identifying trends and patterns, e.g. spatial-temporal representation; (g) annotating, remembering, re-visiting, and setting aside; and (h) developing a sense of what is in the data – exploring what is there.



Chen Yu

Email: chenyu@indiana.edu

CHEN YU is an Associate Professor in the Department of Psychological and Brain Sciences, Cognitive Science and Computer Science at Indiana University. He received his Ph.D. from University of Rochester in 2004 and joined Indiana University since then. Dr. Yu has received the David Marr Best Paper Award from the Cognitive Science Society, an outstanding early contribution award from the International Society of Infant Studies, and a junior faculty award from Indiana University. His research is supported by NIH, NSF and AFOSR. Dr. Yu has published papers in the fields of cognitive science, developmental psychology, and computer science. He has served on NSF panels and NIH study sections.

Primary research interests

Yu's work is inspired by the concept of embodied cognition and he specializes in cognitive modeling, language acquisition, perception-action coupling, cognitive development, and multimodal human-robot interactions. He studies human development and learning through both behavioral studies and computational modeling. In particular, his research focuses on how language is grounded in sensorimotor experience and how language development depends on complex interactions among brain, body and environment. He is also interested in connecting human behavioral studies with artificial intelligence research in computer science to develop biologically-motivated artificial agents. Most recently, Yu has been analyzing micro-level fine-grained behavioral data, such as eye movement data, body movement data, speech and video data streams.

Challenges for Open Data Sharing

- Exploratory data analysis: how to integrate top-down theoretical knowledge with bottom-up patterns extracted from raw data to effectively discover new findings
- Data management and automatic data annotation: how to find and take advantage of new techniques in computer science
- Data visualization
- Data reduction through multiple stages
- Technical issues in data sharing across labs (IRB and parental consent, in particular with video data)

Researchers



Paulo Carvalho Email: pcarvalh@indiana.edu

Paulo Carvalho is a Ph.D. student in the Department of Psychological and Brain Sciences at Indiana University. Carvalho is interested in the real-time processes that drive what people learn. His current work examines task dynamics during categorization. In children, he is studying how toddlers use comparison and labels to learn about objects. In adults, he is examining how people combine previously learned perceptual information and feedback as they learn new concepts. He is also interested in the implications of these processes for educational practices.



Caitlin Fausey Email: cfausey@indiana.edu

Caitlin Fausey is a Postdoctoral Fellow in the Department of Psychological and Brain Sciences at Indiana University. She received her B.A. from Northwestern University in 2004 and her Ph.D. from Stanford University in 2010. Fausey's research focuses on interactions among language, attention and memory. She takes both cross-linguistic and developmental approaches to examine how differences in language use can lead to differences in what people learn from their experiences. Fausey currently directs projects about mechanisms of early language development, including multisensory integration and dynamics of repetition in word learning.



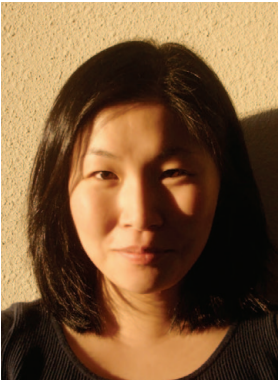
John Franchak Email: franchak@nyu.edu

John Franchak received his PhD in Psychology from New York University in 2011. His dissertation work explored perceptual-motor development across the lifespan, focusing on the real time processes involved in motor control. John has worked to adapt head-mounted eye-tracking methods for use with infants in natural interactions and has developed data coding techniques to cope with natural behavior. John is currently a postdoctoral fellow at New York University researching infants' visual exploration of complex, real-world stimuli.



Damian Fricker Email: dfricker@indiana.edu

Damian Fricker is pursuing a joint PhD in Psychological and Brain Sciences and Cognitive Science at Indiana University Bloomington. His research focuses on Human-Robot, Human-Human, and Parent-Child interactions consisting of large quantities of multi-modal data from eye-trackers, motion-trackers, video, and audio data streams. He has been using an in-house visualization tool integrated with various data mining techniques over the past 2 years to analyze the temporal data streams.



Amy Joh Email: amy.joh@duke.edu

Amy Joh received a Ph.D. in Cognition and Perception from New York University in 2006. In 2008, following postdoctoral fellowships at the University of Massachusetts, Amherst, and University of Virginia, she joined the faculty of Psychology and Neuroscience at Duke University. Amy is interested in the relationship between perceptual-motor skills and cognitive development in infants, children, and adults. The research in her lab focuses on how people gather and use different types of perceptual information for goal-directed actions, how different learning mechanisms influence the development of actions, and the role of cognitive strategies on planning and executing actions. For the past 10 years, she has used MacSHAPA to explore and understand the findings from her behavioral research.



Lana Karasik Email: lana.karasik@nyu.edu

Lana Karasik received her PhD in Developmental Psychology from New York University in 2009. Her research examines relations between infants' growing perceptual-motor skills and social-cognitive abilities over the first two years of life in social and cultural contexts. She has conducted longitudinal and cross-sectional, observational and experimental, and within- and cross-cultural studies of infants and mothers examining the interdependent relations among infants' emerging skills and the social environment. Using MacSHAPA/OpenSHAPA she has developed efficient data coding methods to capture infants' behaviors across multiple domains—locomotion and manual and social action—and changes in mothers' behaviors. Currently, Lana directs two large-scale, longitudinal, school-based studies of children's self-regulation and early academic abilities in Dr. Clancy Blair's lab at New York University.



Celeste Kidd Email: ckidd@bcs.rochester.edu

Celeste Kidd is a doctoral candidate in Brain & Cognitive Sciences at the University of Rochester. She is advised by Richard N. Aslin, with whom she works on research concerning infants' early attention, information processing, learning, and cognitive development. In 2008, she was awarded an NSF Graduate Research Fellowship. In 2010, along with collaborators Steve Piantadosi and Richard Aslin, she was awarded a Computational Modeling Prize for Perception/Action by the Cognitive Science Society for a paper on infants' allocation of attention. Her dissertation research (in-progress) investigates how infants build high-level representations from low-level sensory statistics using a combination of statistical modeling and infant eye-tracking methods. She has been using OpenSHAPA for the past two years to code visual fixations and gestures during infant-parent play interactions. Preliminary results of this work were presented at the Society for Research in Child Development 2011 Biennial Meeting in Montréal, Québec. Celeste received her B.A. in 2007 from the University of Southern California, where she worked on language development, production, and processing topics.



Michele Lobo Email: malobo@udel.edu

Michele Lobo, PT, Ph.D. is a research scientist in the physical therapy department at The University of Delaware. She received her bachelor of science degree with a major in biology and a minor in psychology from The College of New Jersey in 1994. She earned a master of physical therapy degree from The Medical College of Pennsylvania and Hahnemann University in 1997. After several years of pediatric practice, she returned to academia to earn her doctorate in Biomechanics and Movement Science from The University of Delaware in 2006. She has worked as a visiting scholar at the Free University in Amsterdam as well as The University of California in Berkeley. She has been the recipient of multiple fellowships and awards, including the Dorothy Briggs Memorial Scientific Inquiry Award from The American Physical Therapy Association in 2006 and the Lolas E. Halverson National Young Investigator Award from The Motor Development and Learning Academy Committee of The National Association for Sport and Physical Education in 2008.

Dr. Lobo's research program focuses on describing and understanding developmental processes and designing interventions to advance development for individuals with motor and cognitive impairments. She has 4 years of experience using MacSHAPA for behavioral coding. She also has experience collecting and processing data using motion analysis systems. In the past she has written programs using LabView software to process data. She is now a trained Filemaker relational database developer and designs databases using this software to store, process, relate, and export a variety of data types, including the coding data generated using MacSHAPA.



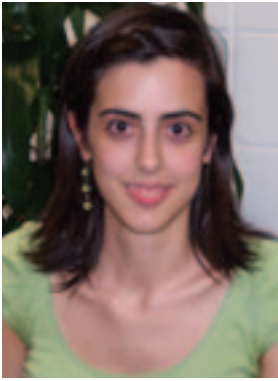
Evan Patton Email: pattoe@rpi.edu

Evan Patton is pursuing a Masters of Cognitive Science and Ph.D. of Computer Science at Rensselaer Polytechnic Institute. His masters research focuses on automated generation of stochastic CPM-GOMS models from log files of human behavior. His Ph.D. work includes collaborative data frameworks using ontologies over the World Wide Web, integration services for environmental data at federal and state levels, and distributed knowledge representation and reasoning on ad-hoc mobile networks.



Kasey Soska Email: kcs5q@virginia.edu

KASEY C. SOSKA is a Post-doctoral Research Associate in the Department of Psychology at the University of Virginia. He received his B.S. from Penn State University in 2005 and his Ph.D. from New York University in 2010. Dr. Soska received dissertation awards from the Society for Research in Child Development and the International Society for Infant Studies. Soska's research documents the developmental relations among action, perception, and cognition. In particular, he investigates how developing motor skills promote learning. He has explored the development of postural control, manual actions, crawling, and walking and their downstream effects on object and motion perception. His methods include visual habituation, preferential looking, desktop and head-mounted eye tracking, event-related potentials (ERPs), and behavioral coding with infants, young children, and adolescents. His research has focused on typically developing children, as well as adolescents with autism. Soska has been using MacSHAPA and OpenSHAPA for the past 7 years and is interested in developing ways for new researchers to produce datasets rich in human behaviors.



Catarina Vales Email: cvaless@indiana.edu

Catarina Vales is a PhD student in the Department of Psychological and Brain Sciences at Indiana University. She is interested in the dynamics of attention and learning, particularly in the context of language acquisition in the first years of life. Her research seeks to understand how attention develops, how children use regularities in the environment to learn new words, and how attentional processes might organize themselves into a coherent mechanism that supports language learning during development.



Tian Xu Email: linger.xt@gmail.com

Tian Xu (Linger) is a student in Computer Science and Cognitive Science joint PhD program at Indiana University Bloomington where she is advised by Dr. Chen Yu. Her main research interest lies in examining how learning is grounded in sensorimotor experience for infants and building computational models to explore joint attention in human-human interaction and human-robot interaction. She has been working on coding and performing temporal data mining on eye gaze fixations, manual actions and speech patterns in child-parent interaction and human-robot interaction.



making research easier
exploratory sequential data analysis