



The Brunswik Society Newsletter

www.brunswik.org

Compiled by Clare Harries (clare.harries@ucl.ac.uk)

October 2003

In the 100th Anniversary of Brunswik's birth the newsletter is full of interesting research reports, theoretical discussions and proposals for future work. Many thanks to all those who have contributed.

Contents:

2	Abdollahi	Effects of Pursuing Useless Information on Decision Making across Cultures
2	Adelman	Evaluating Argument Representations
2	Anderson	Lopsidedness in the Study of Judgment and Decision Making: Analogy & a Prod
3	Araújo, Davids, Serpa	An Ecological Approach to Expert Decision Making in Sailing
4	Athanasou	Factors Influencing Job Choice
5	Brown & Robyn	Reserving a Key Place for Reality: Philosophical Foundations of Theoretical Rotation
5	Cooksey	Hierarchical Lens Models, Think Aloud Protocols & Teacher Judgments
6	Cvetkovich	Assurance, Trust & Advice Taking
7	Dal Santo	Decision Processes & Performance of Mining Engineers
7	Denig	Prescribing Decisions & Information Search Strategies
8	Dhami	Predicting Dangerousness
8	Dibadj	Statistical characteristics of the lens model equation parameters
8	Doherty	The Small Window Effect, Reasoning & Task Environments
9	Dunwoody	Organism-Environment Interactions & Human Cognition
9	Earle	Trust, Confidence and Cooperation
10	Elbogen, Tomkins	Mental Health Clinicians' Judgments About Violence Risk
11	Fasolo, McClelland, Todd	Attribute correlations & weight distributions influence frugality of preferential choice process
11	Gattie & Bisantz	Using the Lens Model for modeling dynamic environments
12	Gunderson	Human Preference Discovery and Robotics
12	Hamm	Public, Patient, & Physician Beliefs about the Relation Between Human Papillomavirus & Cervical Cancer
13	Hammond	Taxonomy of tactics
13	Holzworth	Research in the Brunswikian tradition continues at the University of Connecticut
14	Karlsson & Juslin	Judgment Processes: Ecology-Fostered Representations, Input Samples & Human Limitations
15	Kirlik	The Technological Ecology
15	Klayman	More About Confidence
16	Kostopoulou	Between-Doctor Variation in COPD Gatekeeping Decisions
16	MacCormick	Patient Prioritisation for Elective General Surgery
17	McKiernan, Wolf, Figueredo	Extinction and Habituation in the Context of Brunswikian Learning Theory
18	Mumpower	Brunswikian Approaches to Policy Analysis
18	Newell	Recognition, rolling regression, control & insight
19	Paashuis & Kerstholt	Shoe Print Investigation: The Process & Factors of Influence
19	Paterson, Dowding, Cassells	Nurses' & Psychiatrists' Judgements about Suicide Risk & Decisions on Observation
20	Piperides	Expert Witness Use of Consistency Cues: Judging the Veracity of Testimonies
21	Quesada	Latent Problem Solving Analysis (LPSA): A Computational Theory of Representation in Complex, Dynamic Problem Solving Tasks
23	Rothrock	Use of Multi-objective Optimization toward Development of a Noncompensatory Lens Model
24	Schaefer	Metacognition in Multiple Cue Probability Tasks
25	Schwartz	The Statistical Structure of Speech Sounds Predicts Musical Universals
25	Shanteau & Hilton	Overconfidence vs Self-Confidence in Experts
26	Smith, Walker, Gilhooly	Clinical Guidelines on Depression: A Qualitative Study of GPs' Views
28	Snook & Taylor, Bennell	Brunswik's Influence on Geographic Profiling Research
29	St-Cyr & Vicente	Effects of Data Uncertainty in a Process Control Microworld
30	Stewart	Ecological Analysis & Judgment Analysis
30	Tape	Logistic Lens Model analysis in Medicine
31	Tweney	A Graphical Artifact by Brunswik
31	Weaver	Modeling & Mentoring
32	Wigton	Communication, Judgment & Decision Making in Medicine
32	Wolf	Brunswikian Representative Design in the Research of Persistence in Educational Psychology
33	York & Barclay	Clear Logic and Fuzzy Guidance: A Policy Capturing Study of Merit Raise Decisions
34	Nineteenth Annual Meeting of the Brunswik Society in Vancouver	

Effects of Pursuing Useless Information on Decision Making across Cultures

Abdolhossein Abdollahi

Zarand Islamic Azad University, Iran
a-abdollahi@iaiu-zarand.ac.ir
a_abdollahi@yahoo.com

This year, I investigated the effects of useless information on decision making across three individualist and collectivist cultures (i.e. American, Iranian, and Japanese) using the paradigm introduced by Bastardi and Shafir (1998).

An experiment was conducted in which 144 American, Iranian, and Japanese participants (Nurses in Problem 1 and university undergraduate volunteers in Problem 2) responded to two surveys adapted from Bastardi & Shafir (1998, Problem 1) and Redelmeier, Shafir, & Aujla (2000, Problem 2). Half of the participants received the Simple version and the other half, the Uncertain version. In the Simple version, no "useless" information was provided, but in the Uncertain version participants received information that was noninstrumental for the task of decision making. No incentives were given to the participants.

The results revealed that participants in all three cultures are susceptible to pursuing useless information under uncertainty. Specifically, there was a significant difference between the participants' responses in the Simple and the Uncertain conditions. It was also found that participants from the individualist (American) culture show a higher rate of pursuing the useless information than the other two groups.

The findings of the present experiment demonstrate that people from the American, Iranian, and Japanese cultures tend to use noninstrumental and irrelevant information when making decisions. Also, participants with an individualist background may show a higher level of proneness to the effect. Bem's Self-perception theory may explain this observed difference.

Evaluating Argument Representations

Len Adelman

George Mason University, USA
ladelman@gmu.edu

Colleagues at the MITRE Corporation and I have begun research evaluating various methods for graphically representing arguments for judgment problems without a correct answer. Our first experiment evaluated the argument structure

representation developed by the philosopher Stephen Toulmin. These graphic structures show the claim (including any qualifications about it), the data supporting the claim, the warrant (or rule) for connecting the data to the claim, further backing (or data) supporting the warrant, and rebuttals that undermine the argument. Our experiment examined the effect of using Toulmin argument structures on participants' (a) evaluation of the logical soundness of arguments in two newspaper-length articles varying in their organizational clarity, and (b) communication of the argument only using the Toulmin structure. Initial data analyses suggest that the Toulmin structures were effective for the evaluation, but not for the communication of the argument in the more poorly organized article. There were no effects for the article with the well organized argument, presumably because structuring was not needed for it. Future research will evaluate other argument representations, including variations on a "policy capturing" representation scheme.

Lopsidedness in the Study of Judgment and Decision Making: Analogy and a Prod

Christopher J. Anderson

Temple University, Philadelphia, USA
chris.anderson@temple.edu

Lately I have been thinking about a distinction that is not often made and a resulting lopsidedness in the study of judgment and decision making.

As scholars of judgment and decision making, our primary concern is in how people respond to their environment by selecting actions to perform in that environment (decisions). Judgments are usually thought of as assumptions or hypotheses about the environment that guide decisions.

Now, in considering how to study such phenomena, consider the following analogy to an early human: our subject has various pressing problems of survival to solve, one of which includes procuring food. This constitutes a goal. The environment provides affordances for several methods of reaching this goal. Let us suppose that the individual in question is part of a hunting team that aims to procure meat. In trying to anticipate this person's success, a crucial distinction is salient: one is the selection of tools which the individual uses, and the other is the ability for the person to use that tool to reach the goal. For example, our hunter might be using an appropriate and well designed spear, but be poor at throwing it, and thus be unsuccessful at reaching the goal. Under another unsuccessful condition, the spear might be well designed, but the environment blocks its use – perhaps it is too windy to throw it, or one tries to

throw it in an environment with obstructions. Perhaps one has the ability, and the environment supports its use, but the social context is a problem – the other hunters on the team make mistakes that prevent the goal from being reached.

If one applies this analogy to decision making, then heuristics, knowledge, and learning comprise some of the important tools. And the field is strong in its emphasis on studying this tool box, to the point where disagreements about what is in the box and their effectiveness is the most visible discussion in the field. However, the other components of the system remain less well understood. Here are some consequences of the current distribution of scholarship:

1) Goals: lack of emphasis on understanding the goals of judgment and decision making creates distortions. For example, in studying inaction effects that are due to emotion, it had been commonly been assumed that delays that lead to material loss are due to biased thinking. However, individuals have goals for decision making that go beyond accuracy and material gain; they also wish to cope with negative emotions, and have other goals, such as preserving social identity. Understanding the goals people have for judgment and decision making will further our understanding of decision making processes, and take us beyond merely labeling a judgment/decision strategy as “biased” or “adaptive.”

2) Tool vs. User: In focusing on the knowledge people apply to problems, we tend to miss the role played by the user of the tool and their ability to use a tool effectively. This includes, for example, important issues such as the reliability of judgment/decision strategy use, and the ability to self-regulate enough to carry out a judgment or decision plan effectively. These factors can be just as important in determining whether an individual is successful in reaching goals as the nature of the mental tools they use, yet we know much less about them. Brunswikian researchers may have an advantage in this area in that the lens model makes a distinction between knowledge and consistency, and research has been carried out that focuses largely on user issues, but I believe this work has not yet made an impact proportional to its importance to the actual success of judgment.

3) Role of the environment: I hardly need to lecture a group of Brunswikians on the role that the environment plays. But again, in focusing largely on the tools, the field misses understanding how those tools fit different environments, and what affordances the environment allows in terms of success of the decision or achievement of accuracy, a point that is just beginning to be acknowledged by a wider audience.

In a sense, I have the feeling that I am preaching to the choir about these issues, in that the Brunswikian tradition has been more balanced with respect to these issues than other traditions. However, I feel that it is time that we begin designing programs of research that highlight the lopsidedness of the field and persuade our colleagues that a more inclusive view will increase the value of our work manifold.

An Ecological Approach to Expert Decision Making in Sailing

Duarte Araújo

Faculty of Human Kinetics,
Technical University of Lisbon, Portugal
Daraujo@fmh.utl.pt

Keith Davids

School of Physical Education,
University of Otago, New Zealand

Sidónio Serpa

Faculty of Human Kinetics,
Technical University of Lisbon, Portugal

Following an ecological approach, the primary objective of this study was to examine expertise effects on decision-making in sailing. Dynamic contexts in sailing were provided through interactive computer simulations, used to reveal the utilisation of information, and the active exploration of the environment by participants (n=23 non-sailors and n=35 sailors). Sailors were divided into three groups according to ranking in sailing expertise: *Experts*, *Skilled*, and *International*. *Experts* were world-class sailors, *skilled* participants were regular competitors at international level, and *international* participants had represented the national team at least once in competition. During the simulated regatta a concurrent verbal protocol analysis was used to measure responses to four sources of information: adversary, spatial, manoeuvres and wind. Simultaneously, participants pressed keys in a keyboard to direct the boat, registering two categories of actions: technical actions and adjustment actions. The computer automatically controlled the dynamics of sail and boat balance. Outcome variables, final classification and total time were also recorded.

This dynamic task was based on Brehmer's (1996) concept of microworlds, and incorporates the representative design issues that Hammond (1999) raised about such tasks. To measure cognitive processes *functional* fidelity is of greatest importance, being relevant to use a partial (visual) simulation, like

the one in this study. The validity of this simulator was verified by asking seven national head coaches from Portugal and Holland to rate the correspondence between the competition presented on the simulator and real world competitive events, with respect to on-line decision-making. They were asked whether the simulation would allow sailors to manifest their "typical" decision making behaviour and tactics in competition. Despite the consideration that the simulation could be improved on several features (i.e. physical fidelity), all the experts agreed (100%) that it was an effective method to assess decision-making in competition. Sailors also reported the same level of agreement as the expert coaches after the experiment.

Results verified that expertise level was significantly related to total time and final classification, indicating a positive relationship between level of expertise and performance on a simulated regatta. Statistical analyses showed that non-sailors significantly differed from sailors in the use of adversary and wind information during the regatta. But, there were no significant differences among the sailors' groups. Non-sailors performed significantly more actions than sailors, during most of the regatta. However, polynomial trend analysis revealed that each group of sailors exhibited specific patterns of information utilization and performed actions indicating different trends in performance. Decision-making in sailing is characterized by non-linear accumulated effects of exploiting environmental features and information use in a regatta, which are dependent on the level of attunement to the specific task constraints.

The relative weighting of information utilization and actions performed in this study denotes the probability of use by participants and not that all information is used at the same time with different weightings attached. Depending on the specific constraints, multiple sources of information were available to control task performance (see Savelsbergh & Van der Kamp, 2000). Data showed slight differences in the probability of use of certain information sources and the performance of certain actions in each regatta phase, which accumulated from phase to phase, leading to robust differences among groups. These findings emphasized why development of an adequate theory of how expertise effects in decision-making are generated, is dependent on an accurate account of exactly what those effects are because the latter can put strong constraints on the former (Vicente & Wang, 1998).

Brehmer, B. (1996). Man as a stabiliser of systems: From static snapshots of judgment processes to dynamic decision making. *Thinking and Reasoning*, 2, 225-238.

Hammond, K. R. (1999). Mats Bjorkman and the Swedish studies of judgment and decision making. In P. Juslin, & H. Montgomery (Ed.), *Judgment and decision making: Neo-Brunswikian and process-tracing approaches* (pp. 305-320). Mahwah, NJ: LEA.

Savelsbergh, G., & Van der Kamp, J. (2000). Information in learning to coordinate and control movement: is there a need for specificity of practice? *International Journal of Sport Psychology*, 31, 467-484.

Vicente, K., & Wang, J. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review*, 105, 33-57.

Factors Influencing Job Choice

James A. Athanasou

Faculty of Education,
University of Technology, Sydney, Australia
Jim.Athanasou@uts.edu.au

This research sets out a Perceptual-Judgemental-Reinforcement approach to job choice under conditions of complexity and uncertainty. It investigates the claim that job choices are based on six implicit factors: such as the specific size of the occupation, the proportion of employees working full-time, the earnings, the job prospects, gender dominance in an occupation, the level of unemployment in the occupation and the predominant age group in the job.

Nine case studies involving choices from 25 randomly selected advertised jobs are presented. Results indicated substantial idiosyncrasy in job choices.

This study used the single lens model of analysis. Job choice was the dependent binary variable and independent variables were the seven implicit labour market factors. An individual logistic regression indicated no statistically significant influence of key labour market indicators in any of the nine case studies.

It was concluded that job choice was idiosyncratic; that individuals lacked insight into their job choices and probably relied upon relatively few unstated cues. The findings have direct implications for the relevance of occupational information and for key issues in the delivery of vocational guidance.

Reserving a Key Place for Reality: Philosophical Foundations of Theoretical Rotation

Steven R. Brown and Richard Robyn

Department of Political Science

Kent State University, USA

sbrown@kent.edu, rrobyn@kent.edu

Factor rotation has been a controversial topic in the history of factor analysis, and preference has always been for solutions that are determinant, such as the simple-structure solution approximated by varimax. William Stephenson's preference for judgmental rotation, available in Q methodology software packages such as PQMethod and PCQ, is little used, due in part to lack of understanding of its philosophical bases in the writings of Egon Brunswik (psychological cues), J.R. Kantor (specificity), Charles S. Peirce (abductive logic), and Michael Polanyi (tacit knowledge).

Brunswik emphasized the ways in which cues enable organisms to move about in the external world in the tacit ways that Polanyi described, and he was critical of "systematic" designs that glossed over the probabilistic relationship between cue and goal. As is known in Q methodology, which Stephenson invented, individual statements can take on different meanings depending on the context in which they are embedded, and so background cues are required for orientation; he therefore incorporated Brunswik's idea that the stimulus environment had to be sampled (as in Q samples) so as to provide the Q sorter with a variety of cues representative of the stimulus situation.

Stephenson also incorporated Brunswik's conception of cues as applied to factor rotation: "... if persons A, B, C, and D can be shown, sociometrically, to be linked to another, E, the investigator may have a 'hunch' that factor solutions centered upon E, rather than upon A, B, C, or D, will prove pregnant in some way. There are countless 'cues' or 'tricks of the trade' of this kind in every science. All are deliberate impositions of inferential and empirical possibilities upon otherwise neutral situations." (This quote appears in Stephenson's *The Study of Behavior*, University of Chicago Press, 1953, p. 39, for which Brunswik served as external reviewer.) That is, theoretical rotation permits features of reality (as understood by the factor analyst), and not simply statistical principles, to play a role in the final factor solution.

The principles of theoretical rotation are illustrated in studies of French national identity, trait analysis, and reader response to fiction, and the conclusion is reached that statistical solutions, such as varimax, respond only to the surface features of data, but do not necessarily penetrate to their operant core.

(Paper presented at the 19th annual conference of the International Society for the Scientific Study of Subjectivity, Kent State University-Stark Campus, Canton, OH, 2-4 October 2003.)

Hierarchical Lens Models, Think Aloud Protocols and Teacher Judgments

Ray W. Cooksey

New England Business School

University of New England, Australia

rcooksey@pobox.une.edu.au

I am re-surfacing this year after a couple of intensively administrative years here at UNE. My report provides an update on a project I initially reported on in the 2001 Brunswik Newsletter. The project was an Australian Commonwealth Government-funded collaborative research endeavour with colleagues from Griffith University in Queensland. We employed judgment analysis, in conjunction with a think aloud protocol methodology, to explore teachers' judgments of Year 5 student achievement in writing tasks. The study investigated how 20 teachers made judgments of writing achievement, when confronted with specific exemplars of students' written work in particular genres. Teachers worked through a series of 25 pieces written by their own students (judgments made 'in the classroom context') and a series of 25 pieces sampled (by the researchers) from a wider group of students studying with other teachers (judgments made 'out of the classroom context'). Thus, the judgment ecology was partly tailored to each teacher. We predicted that teachers would import and apply more personal knowledge about the task and the writer, when they knew who wrote the piece they were assessing than they would if the author of the work was unknown to them. Two passes were made through the pieces of writing by each teacher: first making assessment judgments as they normally would in their class and then making judgments using the benchmark assessment framework created by the Queensland Department of Education. Teachers were asked to think aloud as they made each of their judgments. The objective here was to compare classroom assessment practices with a system-wide mandated benchmarking framework. Think-aloud protocols were exhaustively coded and condensed into four central cues for judgment: positive and negative comments about the writing context and task and about the student writer. All pieces of writing were separately coded in terms of objectively observable text and writing features; cue condensation resulted in a set of 5 coded text features. Idiographic hierarchical judgment regression models were

employed, in a hierarchical lens modelling framework, to estimate the respective contributions of judgment context, think-aloud coded features, coded text features, and interactions between coded think-aloud features and judgment context. The lens model framework allowed us to compare assessment and benchmark policies within teacher (judgment agreement ranged from .49 to .88). For most teachers, coded think aloud cues (especially positive and negative task-related comments) and two of the five coded text-related features (length and complexity of the writing and adherence to the formatting requirements and conventions of the genre) contributed to judgment, for both assessment and benchmark judgments. However, the use of student-related think aloud information was strong for some teachers, but not others, and was more prevalent for assessment judgments than for benchmark judgments. A small number of teachers revealed substantive differential reliance on think aloud information for their 'in context' and 'out-of-context' judgments – somewhat consistent with our prediction, but not a generalisable phenomenon. For most teachers, assessment judgments were more consistent than system-mandated benchmark judgments. One implication seems to be that system-mandated judgment policies are more problematic for teachers to apply not only for their own students' work, but also when they are asked to make decontextualised judgments (as is often demanded in education systems where teachers mark and/or moderate the assessments of students they do not teach). Another implication of the research is that judgment analysis and think-aloud protocol methods can be successfully employed, in a synergistic fashion, to enhance insights into judgment processes. We will soon be undertaking a more in-depth qualitative analysis of the think aloud protocols to see if we can map case-by-case configural judgment dynamics – a time consuming task!

Assurance, Trust and Advice Taking

George Cvetkovich

Western Washington University, USA

George.Cvetkovich@wwu.edu

I have continued to investigate the information processes involved in judgments to rely on another person. Three lines of evidence -- 1) studies of social motives and social factors influencing behavior and experience, 2) studies of cognitive functioning and information processing, and 3) functional brain imaging studies -- indicate that it is useful to conceptualize the two routes of reliance shown in Table 1 -- the route of assurance and the route of trust. Assurance is based on primary inference using immediately available information (proximal cues) defining situational characteristics. Situational characteristic affecting willingness to rely on another person include systematically enforced laws, procedures attempting to ensure fair and just decisions, institutionalized accountability of risk managers, a good performance record, and opportunities to voice one's view. Accumulating evidence indicates that the identification of situational characteristics, including the presence of assurances, occurs through the activation of the ventral information processing stream of the brain. It has been called the "what" system of information processing because it functions in the recognition of objects, individuals, events and situations.

Reliance based on the route of trust involves the making of secondary inferences of attributions about other individuals (distal states). Attributions about others -- what their motives, goals, morals, personal characteristics are, in essence, who they are -- are based on social representations of our knowledge of a particular individual or members of particular group, as well as representations of how the human mind works in general. The colorful term "mind reading" has been applied to this process to convey its important function. Mind reading is basic to many important human interactions. Efforts to mind read,

Table 1. Routes of Reliance

Type of reliance	Information used	Information Processing	Mode of processing	Judgment
Reliance based on assurances	Social representations of situation (e.g., social relationships, laws, contracts, voice, justice procedures)	Ventral (what) stream	Combination of implicit (emotion) & explicit (conscious)	Rely or not rely
Reliance based on trust	Social representations of an individual (attributions of morality, personality, motives, goals, etc.); members of social groups (e.g., politicians, lawyers) and/or humans in general	Ventral + Dorsal (where) streams		

including attempting to determine if an individual is trustworthy, activate other parts of the brain in addition to the ventral stream of information processing. The dorsal stream of information processing that runs along the top of the brain, linked to the ventral stream through the temporo-parietal junction, is also activated. The dorsal stream has been called the "where" system of information processing because it functions in the recognition and generation of goal directed actions. Imaging studies show activation in areas of this stream when we are monitoring the mental states of other people and ourselves. Both assurance and trust involve combinations of implicit and explicit processing.

Using this conceptualization (Cvetkovich & Winter, forthcoming), I have continued cooperative research with Pat Winter (USDA Forest Service) on various aspects of National Forest management including protection of threatened and endangered species and wildfires. Preliminary results of surveys of representative samples of residents of Arizona, California, Colorado and New Mexico confirm findings of earlier small sample studies (Cvetkovich & Winter, 2003). Trust is related to perceptions that the Forest Service shares the individual's salient values of forest management. The most salient value for most individuals is protection of wildlife and habitat (as opposed to human use of the forests). Level of trust also is related to the Forest Service's perceived consistency of following salient values in practice. I am also beginning research on the effects of trust and assurance on use of e-commerce as well as continuing work (intermittently) on a book on trust for a general audience.

Cvetkovich, G. T., & Winter, P. L. (2003). Trust and social representations of the management of threatened and endangered species. *Environment & Behavior*, 35(2), 286-303.

Cvetkovich, G. T., & Winter, P. L. (forthcoming). The what, how, and when of social reliance and cooperative risk management. In M. Siegrist & R. E. Löfstedt (Eds.), *Trust, Technology, and Society: Studies in Cooperative Risk Management*. London: Earthscan.

Decision Processes and Performance of Mining Engineers

Lisa Dal Santo

The University of Queensland, Australia
lisa@psy.uq.edu.au

My doctoral research conducted over the past few years has focused on mining engineers in their role of designing structural support for underground mine areas. In this area of "ground control", mining

engineers are required to assess the amount of risk in a situation, and decide if a particular area needs additional support over and above the standard amount. The overall aim of a set of four studies was to arrive at recommendations for training, decision support, or system design for engineers involved in ground control design within the mining industry.

The first two studies were conducted largely in order to explore the domain and to elicit key cues from experienced mining engineers. These studies involved interviews, questionnaires and cognitive task analytic procedures. The third study combined Social Judgment Theory (SJT) and Signal Detection Theory (SDT) methodologies in order to examine risk judgments and support decisions. These techniques provided information about the engineers' cue weightings, decision threshold and accuracy. This study also investigated relations between these measures and other constructs such as risk propensity, risk perceptions and organisational safety climate. The final study was conducted in order to investigate differences between more and less experienced mining engineers, and to assess the adequacy of existing University-level training in ground control. The results of these studies, and recommendations for training to support and enhance decision-making in this area, are in the process of being written up in my PhD thesis.

Prescribing Decisions and Information Search Strategies

Petra Denig

Department of Clinical Pharmacology
University of Groningen, The Netherlands
p.denig@med.rug.nl

In our department several studies are conducted related to prescribing decisions that do not seem to be in line with evidence-based guideline recommendations. In the past, we have tried to capture decision policies using mostly hypothetical cases. We also experimented with giving policy feedback to general practitioners, showing that this can improve their future decision making. In our search for other instruments that may help practitioners to increase the quality of prescribing, a new project will start focussing on the information needs of doctors when deciding on possible treatment during consultation. Special attention will be given to the type of drug information needed in relation to the information cues provided in a representative set of patient cases, and the way doctors search for this information in readily available information sources such as the paper and electronic version of the Dutch Pharmacotherapy Reference

Book. This should help developers of electronic versions of such reference books to increase the value and use of this new information source.

Predicting Dangerousness

Mandeep K. Dhmi

Department of Psychology
University of Victoria, Canada
mkdhami@uvic.edu

Over the past year, I have been working with Grant Broad (gbroad@invicta-tr.sthames.nhs.uk) and Peter Ayton (P.Ayton@city.ac.uk) on a lens model study of forensic psychiatrists' and psychologists' judgments of violence. Grant is an experienced practitioner and is currently conducting a masters' thesis on this issue under the supervision of Peter at City University, London.

The Kent Forensic Psychiatry Service in the UK is frequently asked to make risk assessments on patients by probation, prison, police, and community mental health teams. The question they ask is: Are patients at 'high' or 'low' risk of behaving violently? Currently, data is collected on numerous factors such as the patient's mental health, history of violence, and social ties, and are presented to a team of psychiatrists and psychologists. This team meets, sometimes for several hours, to make a risk judgment. The aim is to make the community safer and provide appropriate care to the patient (which may include secure custody if a patient is judged to be 'high' risk). The risk judgment therefore, has significant consequences for the patient and society. Unfortunately, as past research often suggests, 'dangerousness' is difficult to predict.

The aim of the study was to: (1) Model the forensic team's risk judgments of violent offending using a logistic regression model and the Matching Heuristic (a simple non-compensatory model) to determine which best predicts the team's judgments. (2) Model the environment (i.e., predict violent offending) using both models to determine which best predicts the criterion. And, (3) Compare the models of the forensic team with the models of the environment to measure the team's accuracy in predicting violent offending. We were particularly interested in examining the complexity of the decision making strategies used by the forensic team and the complexity of the environment which they are attempting to predict. We are looking forward to presenting the results of this study at the upcoming Brunswik meeting.

I'm quite excited about Grant's thesis. Not only does this study extend my previous research on the use of

simple heuristics in the legal and medical domains to the clinical domain (e.g., see Dhmi, 2003), but importantly, it represents the first full lens model analysis using the Matching Heuristic.

Dhmi, M. K. (2003). Psychological models of professional decision-making. *Psychological Science*, 14, 175-180.

Statistical Characteristics of the Lens Model Equation Parameters

Hassan Dibadj

University at Albany, SUNY, USA
hd7554@albany.edu

My research is about the statistical characteristics of the parameters of the lens model equation. More specifically, I am investigating the estimation procedures, significance tests and probable biases for Rs [measure of cognitive control], G [Matching] and C [Non-linear matching]. The results will clarify under what circumstances we are able to interpret these parameters and whether we can apply the usual statistical procedure for these parameters.

The Small Window Effect, Reasoning and Task Environments

Mike Doherty

Bowling Green State University, USA
mdoher2@mailstore.bgsu.edu

I do not have much to report, given retirement. I have a chapter with Ryan Tweney in K.I. Manktelow and M.C. Chung (eds.), *Psychology of Reasoning: Theoretical and Historical Perspectives*. Hove, Sussex, GB: Psychology Press. The chapter is called Reasoning and Task Environments: The Brunswikian Approach.

I have also been collaborating with Rich Anderson on Kareev's "small window." Our simulations give results between those of Juslin and Kareev. We conclude so far that Kareev was too optimistic and Peter too pessimistic (see Karlsson & Juslin, this newsletter). Our conclusion, based on simulations to date, is that there are limited circumstances in which people infer correlation more validly from small samples than large ones, specifically when the criterion (in a signal detection sense) is set very high. So far, we (Rich Anderson and I) have just a 2003 Cognitive Science meeting presentation on this topic.

Organism-Environment Interactions and Human Cognition

Philip Dunwoody

Mercer University, USA
drdunwoody@yahoo.com

I am currently on the job market and would appreciate any leads on cognitive openings. I have a couple of Brunswikian related projects currently ongoing. All projects focus on how the organism-environment interaction shapes human cognition/judgment. I have included the preliminary abstracts below. All papers are available online at: www.geocities.com/drdunwoody

Consistency, Utility and base rates

Portions of this work will be presented at this year's Brunswik meeting. People often underweight base rate information when compared to a narrow normative standard. Three experiments demonstrate that base rate usage under direct experience is moderated by the consistency as well as the diagnosticity of base rate information, and that this moderating effect extends from a learning environment of direct experience to tests of verbal judgments. Experiment 1 shows that participants use base rate information more when it is consistent than when it is inconsistent. In Experiment 2, this effect was replicated, and transferred to base rate sensitivity in verbal questions posed subsequently. In Experiment 3, participants' use of base rates was once again moderated by its consistency, but this effect was itself moderated by the diagnosticity of base rate information. These studies demonstrate that base rate usage can be an adaptive response to environmental contingencies. (1)

Introspection and Judgment

Many studies report that people lack introspective access to the causes of their behavior and that introspecting before making a judgment actually hurts judgment performance. Conversely, other research emphasizes that introspection aids in self-knowledge. We review this body of research with the goal of developing a unifying framework. We propose a modular theory of judgment based on explicit recognition of two broad types of introspection and two types of criteria. We draw on current work in cognitive neuroscience and explain how the proposed model resolves the current paradoxical findings. Finally, we discuss further implications for the proposed model. (2)

Cognitive psychology's negative bias

Research in cognitive psychology, including social cognition, often emphasizes negative and mechanistic aspects of human behavior. Some examples of this emphasis include biased attributions, an inability to rationally integrate

information, an inability to understand the causes of our own behavior, and the detrimental effects of affect on information processing. We review examples that emphasize a negative aspect of behavior and cognition, show how they have been spun in a more positive light, and discuss possible reasons for this negative bias. We discuss potential causes of this negative bias in terms of attribution processes, our implicit view of humanity, and an overly mechanistic orientation. (3)

- (1) Dunwoody, P. T., Goodie, A. S., & Mahan, R. P. The use of base rate information as a function of experienced consistency and utility. *Manuscript under review.*
- (2) Dunwoody, P. T., & Martin, L. L., To analyze or not to analyze: A review of how introspection affects judgment. *Manuscript in progress.*
- (3) Dunwoody, P. T., & Martin, L. L., Is cognitive psychology guilty of negative bias and the fundamental attribution error? *Manuscript in progress.*

Trust, Confidence and Cooperation

Timothy C. Earle

Western Washington University, USA
timearle@cc.wvu.edu

Over the last couple of years, in collaboration with Michael Siegrist and Heinz Gutscher of the University of Zürich, I have worked to elaborate and test a theory of trust that we now call the TCC model (for Trust, Confidence, and Cooperation). The TCC model is based on two distinctions. The categorical distinction between agent and object provides the basis for distinguishing between trust (the target of which is an agent) and confidence (in which the target is an object). (The Brunswikian distinction between coherence and correspondence is another useful way of talking about the difference between trust and confidence.) The distinction in degrees between freedom and restriction provides the basis for distinguishing among trust relations that vary in degrees of inclusivity, broadness of scope, and porosity of boundaries. General trust is the freest form of trust; social trust within an exclusive, narrow, closed group is the most restricted. Confidence also varies in degrees of freedom, as expressed in degrees of objectivity. Universal law is the freest form of confidence; personal habit is the most restricted. All forms of trust and confidence that have been studied in the empirical literature can be located within this scheme and related to one another. By providing a simple, summary scheme, the TCC model makes available a basis for cross-disciplinary integration and exchange of knowledge.

The most distinctive contribution of the TCC model is that it provides a general account of the relation between trust and confidence. The relation between trust and confidence is the relation between a community and its means of predicting and controlling behavior. Predictions and controls are accepted and justified only within communities. In general, members of a community endorse predictions that seem to support their favored way of life; they discount and oppose predictions that seem to undermine that way of life. According to this model, confidence, and a stable, happy life, can only be achieved through trust. This point is broadly important both theoretically and practically. On the practical side, consider the following representative situation. A person claims to be an environmentalist and to be opposed to plans to drill for oil in the Arctic National Wildlife Refuge. In discussions with proponents of drilling, the environmentalist is presented with extensive scientific studies, conducted by well-qualified experts, which claim to show that the wildlife in the Refuge will not be harmed by drilling; in fact, the wildlife will benefit. In the face of all this science and logic, the environmentalist, not surprisingly, is unmoved. The scientist's arguments did not increase the environmentalist's confidence. Why? Opposition to drilling, in this case, defines the environmentalist's community. Thus, the association between a pro-drilling conclusion and scientific arguments (which, all else equal, should lead to increased confidence) negates the logic of the latter. In this case, the scientific arguments would be confidence enhancing only within groups that are either pro-drilling or undefined by attitudes toward drilling. As the TCC model shows, and as recent studies have confirmed, confidence presupposes trust, a community of recognition based on shared values.

Mental Health Clinicians' Judgments About Violence Risk

Eric B. Elbogen

Duke University Medical Center, USA
eric.elbogen@duke.edu

Alan J. Tomkins

University of Nebraska Public Policy Center, USA,
atomkins@nebraska.edu

Our research was an effort to examine what clinicians actually *do* when they make assessments of violence risk as opposed to what the social scientific researchers urge mental health practitioners to do. There is an abundant literature indicating what clinicians *ought to do* (the recent MacArthur Foundation Violence Risk Assessment Study is the most prominent and extensive attempt to determine

the empirically-validated correlates of violent behavior). Do these research efforts make a difference in the real life of clinicians? Is there a schism between research and practice?

A team of researchers at the University of Nebraska spearheaded by University of Nebraska-Lincoln Law/Psychology Program (then) graduate student Eric Elbogen under the supervision of Professor Alan Tomkins, spent several years examining clinicians' practices comport with what could be described as "best practices" identified in light of the research. We also looked at other, related violence risk assessment and management issues. We used a variety of techniques in our research efforts, including some Brunswikian approaches for identifying variables and capturing decision making.

Elbogen, E. B., Mercado, C., Tomkins, A. J., & Scalora, M. J. (2002). Perceived relevance of factors for violence risk assessment: A survey of clinicians. *International Journal of Forensic Mental Health, 1*, 37-48.

Elbogen, E. B., Tomkins, A. J., Pothuloori, A., & Scalora, M. J. (2003). Documentation of violence risk factors in psychiatric facilities: An empirical examination. *Journal of the American Academy of Psychiatry and the Law, 31*, 58-64.

Elbogen, E. B., Williams, A., Kim, D., Tomkins, A. J., & Scalora, M. J. (2001). Gender and perceptions of dangerousness in civil psychiatric patients. *Legal and Criminological Psychology, 6*, 215-228.

Mercado, C., Elbogen, E. B., Scalora, M. J., & Tomkins, A. J. (2001). Judgments of dangerousness: Are sex offenders assessed differently than civil psychiatric patients? *Psychiatry, Psychology, & Law, 8*, 146-153.

Elbogen, E. B., Mercado, C., Tomkins, A. J., & Scalora, M. J. (2001). Clinical practice and violence risk assessment: Availability of MacArthur Risk Factors. In D. Farrington, C. R. Hollin, and M. McMurrin (Eds.) *Sex and violence: The psychology of crimes and risk assessment* (pp. 38-55). New York: Routledge.

Attribute Correlations and Weight Distributions Influence Frugality of Preferential Choice Process

Barbara Fasolo

Max Planck Institute for Human Development,
Germany
fasolo@mpib-berlin.mpg.de

Gary H. McClelland

University of Colorado at Boulder, USA
gary.mcclelland@colorado.edu

Peter M. Todd

Max Planck Institute for Human Development,
Germany
ptodd@mpib-berlin.mpg.de

Frugal algorithms that use less than full information, like Take The Best and the Recognition Heuristic (Gigerenzer, Todd and the ABC group, 1999), have been shown to perform as well as more information-intensive algorithms across a large number of inferential domains when the information environment is appropriately structured. In this work we extend the investigation of the efficiency of frugal algorithms to the domain of preferential choice, where the task is to choose the best from a set of options described on a number of attributes about which decision makers might care equally or unequally.

We have conducted a series of simulations testing two aspects of the choice environment assumed to affect the efficiency of a frugal preferential choice process. The first aspect, external to the decision maker and characteristic of a given choice task, is the average interattribute correlation. We constructed hypothetical choice sets characterized by a positive or negative average correlation, with the latter implying trade-offs and therefore signifying particularly unfriendly choice environments (Shanteau and Thomas, 2001). In addition, we considered a real choice set of digital cameras extracted from the top recommendations section of the popular website Epinions.com. The second aspect, an internal characteristic of a given decision maker, is the distribution of the subjective attribute weights, ranging from the case where all attributes are equally important to the unequal case where attributes can be ranked from highest to lowest in order of importance.

Our simulations show two important results: First, regardless of the correlations and weight distributions, the fewer the attributes a chooser examines, the fewer non-dominated options there will be to choose among. Because the presence of non-dominated options makes choice more difficult (Tversky and Shafir, 1992; Dhar 1996), this result indicates the psychological benefits of employing a frugal process that limits information use. It can also

explain previous results depicting choosers as being less satisfied and confident when making choices described by many (as compared with few) attributes (e.g., Jacoby, Speller and Berning, 1974; Malhotra, 1982). Second, the costs (measured as proportion of value lost) of using less than full information are rather small, provided that weights are unequal or correlations are positive. Overall, this work shows that frugal algorithms can work well in the preferential choice domain if the environment is appropriately structured, that is when the decision maker cares unequally about the attributes, or attributes are positively correlated.

Using the Lens Model for modeling dynamic environments

Gordon J. Gattie

University at Buffalo,
The State University of New York, USA
ggattie@acsu.buffalo.edu

Ann M. Bisantz

University at Buffalo,
The State University of New York, USA
bisantz@eng.buffalo.edu

The focus of decision making research has been shifting from static to dynamic environments during the last twenty years. Dynamic environments include not only changes in environmental cue values, but in relationships between cues and environmental states. Previous research has demonstrated decision makers are insensitive to statistical properties of the environment, and may encounter difficulty in determining changes in the underlying probabilistic structure of the environment.

One research project, conducted with Natalia Mazaeva from the University at Buffalo, examined the impact of cognitive feedback on detecting environmental changes. Eighty participants attempted to predict the price of a single share of common stock for 150 fictitious companies using four financial indicators. Participants received one of four possible feedback types: feedback based on all trials completed up to current trial, feedback based on the previous ten trials, both cumulative and recent feedback, or no detailed feedback. Initial results suggest groups receiving some form of cognitive feedback adapted to changing environments better than groups not receiving cognitive feedback.

Participants provided with cumulative feedback scored higher than other groups, and also had the largest increase in achievement from the first half to the second half of the experiment. Interestingly, both cognitive control and linear knowledge for most

participant groups decreased overall; we plan on analyzing nonlinear knowledge in the coming months.

The second project involves modeling general aviation (GA) pilots' behavior and weather severity. Data was collected by Kara Latorella for the Convective Weather Sources Experiment conducted at NASA's Langley Research Center. GA pilots received information about existing weather conditions during the flight as either aural information only, aural information and "out-the-window" cues, or aural information and the Aviation Weather Information (AWIN) display. We are currently conducting semi-structured interviews with subject matter experts and examining performance data to determine salient cues and precise cue values for modeling. Information taken from the preflight and inflight weather situation awareness questionnaire will be used to generate decision policies for all participants. Objective data, such as inflight GPS location, NEXRAD, and independent pilot observations will be used to generate environmental information.

Human Preference Discovery and Robotics

Louise Gunderson

Gamma Two, Denver, Colorado, USA
lgunders@palm.com

For the past few years, I have been working on a new method for discovering the target preferences of criminal populations. The events caused by criminals with similar preferences are modeled as the events caused by agents. It was assumed that the agents were selecting targets using the lens model. Since the targets that they selected were, in their estimation, the "best" targets, it was possible to construct a method to discover those preferences. This method, cluster specific salience discovery (CSSD), was used to discover the preferences of the agents. These preferences were used to create a predictive model of their behavior. A predictive model was created for two types of crime in the city of Richmond, Virginia. The CSSD methodology was used on robbery and breaking and entering crime in 1996. The results of this were used to create a predictive model for 1997 for each type of crime. These predictive models provided a statistically significant improvement over other methods. For breaking and entering this model provided a deviance reduction of 343 on 8 degrees of freedom. For robbery this model provided a deviance reduction of 235.9 on 22 degrees of freedom.

This work was done as a graduate student at the University of Virginia. Having completed this work, and received my doctorate, my husband and I have

moved back to Denver, Colorado. There we have formed a small business (Gamma Two) which specializes in data mining and robotics. We are now working on the incorporation of the lens model into robotic software. More details can be found at www.gamma-two.com.

Public, Patient, and Physician Beliefs about the Relation Between Human Papillomavirus and Cervical Cancer.

Robert M. Hamm

University of Oklahoma
Health Sciences Center, USA
robert-hamm@ouhsc.edu

We are producing and analyzing a data set of patients' and clinicians' beliefs (conditional probabilities) concerning human papillomavirus and cervical cancer. Our work is supported by the STD division of the US CDC.

The human papillomavirus (HPV) has over 100 DNA-defined varieties. At least 18 DNA types are sexually transmitted, and 13 are associated with female cervical cancer. By some counts, HPV is the most common sexually transmitted disease (STD), yet it is not as familiar to the public as the most lethal (HIV) or the symptomatic (gonorrhea, syphilis, chlamydia, trichomoniasis) STDs. Recently FDA approved tests for high risk HPV DNA are useful as an adjunct for Pap smears. In explaining these results, physicians must now talk about HPV (rather than just talking about cervical dysplasia). What do and should they say? What information is retained?

We asked people to state their probabilities for some events on the path from the sex act to cervical cancer. Participants were 160 members of the general public, 43 women who had received a diagnosis of high risk HPV by a DNA test, and 33 health care professionals (17 physicians, 12 nurses, 3 pathologists, and a cervical cancer counsellor). We figure the general public has not been informed by a physician about HPV, the patients have been, and the physicians are the source of the information. We want to compare the knowledge of these groups.

We are presenting these data this fall at the Society for Medical Decision Making and Psychonomic Society conferences. Interesting high order concepts can be derived from the probability responses, such as subjective treatment efficacy (absolute risk reduction). It was not particularly surprising that the patient and public probabilities were highly variable, and the mean for most of the events was higher than the mean clinician probability. It was more surprising that the clinician probabilities and efficacies also were highly variable. Analysis of these answers can

identify the areas where patients need to be carefully educated. It also can give guidance on the areas where clinicians need either to do more research, or to think hard about the questions and get their story straight.

Taxonomy of tactics

K.R.Hammond

krhammond@earthlink.net

My professional life has been somewhat uneven during the past year so I haven't made as much progress on my book manuscript as I had hoped (sigh, 'twas ever thus). There is one topic that I did make some progress on, however. I would like to mention it now because I believe it has become timely, and that is the matter of the relation between the concepts of correspondence and coherence and those of intuition and analysis.

In my manuscript I separate these in terms of "strategies" and "tactics". I was prompted to do this as a result of a chance encounter with a remark by Brunswik in his 1956 book to the effect that psychologists were paying too much attention to the "tactics" of judgment/perception. This separation, as well as a greater emphasis on strategies, leads to a cleaner distinction between these ideas.

There are three main tactics in the judgment process; intuition, quasirationality and analysis. All three can be and are employed within each of the two strategies of correspondence and coherence, which provides us with a 6-fold scheme; three tactics within each strategy. Thus, one can employ the correspondence strategy in an intuitive manner, a quasirational manner, and an analytical manner. Similarly with the other combinations. One nice feature of this scheme is that it requires the clear definition of each term, and the need for examples of the six instances. This arrangement should clear up the relation between all members of this conceptual framework.

This 6-fold framework applies while maintaining the concepts of the cognitive and task continua, a feature that has become more important now that my table distinguishing the tactics of analysis and intuition has become adopted (without acknowledgment) not only by Epstein but Kahneman and Frederick, as well as Paul Slovic in a 2002 publication. I will discuss the significance of this matter under the heading of "mindless merging" in my talk at the Vancouver meeting.

Research in the Brunswikian tradition continues at the University of Connecticut

Jim Holzworth

Storrs, Connecticut, USA
holz@uconn.edu

We continue investigating cognitive continuum theory (CCT). My students and I are continuing a collection of biographical data (biodata) from university students in an attempt to relate it to styles of inductive reasoning. Dennis Thomas (MA thesis) examined Taylor and Tversky's (1996) work on spatial perspective within the frameworks of cognitive continuum theory (CCT) and Cognitive-Experiential Self Theory (CEST). Undergraduate students viewed maps of a town, convention center, and amusement park then described them from memory. Descriptions were classified as route, survey or mixed perspectives. Accuracy was examined in relation to perspective change across maps. Results indicate a mixed, rather than route or survey, perspective is the dominant perspective. Map tasks induced changes in mode of cognition, supporting CCT. Very little support was found concerning a relationship between spatial perspective and cognitive style (CEST).

Kathlea Vaughn (MA thesis) investigated the effect of distribution type on decision accuracy in sample-size tasks, based on work by Sedmeier and Gigerenzer (1997). Six sample-size tasks were completed by 202 undergraduate students. Results indicated that participants in the frequency distribution condition attended to sample size significantly more than participants in the sampling distribution condition. Multiple regression analysis showed that individual differences in thinking style, as measured by Epstein's rational-experiential index, accounted for a significant proportion of variance in task performance beyond that explained by distribution type.

Liz Kramer used multiple individual difference variables as screening measures for a serially presented, visual discrimination task. Individuals were categorized as having an analytic or holistic cognitive style, high or low short-term memory (STM) capacity, and high or low levels of rationality, experientiality, and adaptive decision-making styles. Participants received easy or difficult training for the serially presented discrimination task, and then transferred to novel discriminations. Training content interacted with cognitive style and STM capacity to affect transfer accuracy performance. Individual differences in decision-making styles, measured by Epstein's rational-experiential index, did not affect transfer performance.

Kris Canali and I continue our research comparing methods of judgment analysis (JA). Along with five JA methods (OLS regression, ridge regression, smart

ridge regression, subjective weighting, and equal cue weighting), efficient and representative designs are compared. Efficient designs are associated with more accountable judgment variance than representative designs, without significant loss in cross-validations. With fewer cases to judge and not compromising reliability of findings, efficient designs save time and may be more widely used in organizations.

Judgment Processes: Ecology-Fostered Representations, Input Samples and Human Limitations

Linnea Karlsson

Umeå University, Sweden
linnea.karlsson@psy.umu.se

Peter Juslin

Uppsala University, Sweden
Peter.Juslin@psyk.uu.se

Applying a cognitive science perspective to multiple-cue probability learning research and to probability judgment is what signifies our current efforts and it opens several issues worthy of investigation. Under development is an attempt to describe human judgment processes and their capacity constraints. In several experiments we have shown that the structure of the underlying cue-criterion relations is a powerful predictor of what cognitive process and representation a judge will resort to when making judgments. In probabilistic tasks where the cue-criterion relations are best described by an additive function, people are able to abstract these underlying rules and are also able to use these rules when making judgments. However, as soon as the underlying structure is best described by a multiplicative function people have to resort to their specific memory for exemplars. This observation is interpreted as a capacity constraint on controlled, analytical thought. We are constrained to engage in linear and additive combination of information. This constraint implies that it is only in tasks where the cues combine in an additive and linear manner – or tasks well approximated by such a function – that we can rely on explicitly abstracted cue-criterion relations; in other tasks we have to resort to exemplar memory.

Trying to explain why linear and additive models traditionally are seen to describe judgment data quite well, is continued in a series of experiments where we will use continuous instead of binary cues. This will put our research more in line with common multiple-cue judgment tasks and will be more suitable for lens-model analysis. Still we hope to be able to apply our exemplar/cue abstraction paradigm to such

circumstances, which obviously show an increase in complexity.

Together with Ben Newell we are also having another look at insight into multiple cue judgment. One of our concerns is with why multiple regression has become the "golden standard" for assessing insight into judgment. On some assumptions about the cognitive process it might be meaningful to elicit insight into the process in terms of people's insight into the (Beta) weights of the individual cues, but not for all cognitive processes. After all, the linear additive equation is just one way of modelling the environment and people may sometimes be well justified in using quite other processes where the abstracted weight of each cue plays a minor role. We are also trying to define a "common currency" for comparing different ways of eliciting insight into a) the task and the b) cognitive process.

We are further pursuing some issues in research on confidence and probability judgment. We have developed a naïve sampling model to explain why confidence judgments are well calibrated with some probability assessment formats (e.g., the half-range format), yet extremely overconfident with other formats (i.e., interval formats). The idea is that people make probability assessments by retrieving similar exemplars from memory and reporting the observed statistics in these samples without corrections. When this simple idea is combined with different assessment formats the implication is indeed contradictory conclusions about over/underconfidence bias depending on the format. We are further running a series of learning experiments that aim to investigate what sort of processes and representations that underlie probability judgments (e.g., frequencies, representativeness, exemplars), and to ascertain the conditions that promote the one or the other process over the other.

In a recent article (Juslin & Olsson, in press) we have scrutinized Kareev's interesting argument to the point that a limited working memory (sample size) may be an adaptive benefit for early detection of covariation in the environment. We argue that the analysis of adaptive benefit of different sample sizes needs to be performed in terms of the posterior probabilities of the various outcomes, something that implies taking not only Hits (correct detection of useful correlations) but also other possible outcomes (e.g., False alarms) into account. With this kind of analysis we find no support for an adaptive advantage of small sample size and the detection of correlations generally improves with sample size. We propose that the circumstances for the alleged benefit of small samples need to be more carefully specified and demonstrated. Our comment

is forthcoming in the Psychological Review. (See also Doherty, this Newsletter).

Finally, Klaus Fiedler and Peter Juslin are currently putting together a book volume. One of the themes of the volume is to explore the possibilities of placing more of the explanatory burden in regard to judgment phenomena in the input samples to which cognitive processes are applied, rather than by assuming heuristic cognitive processes that conflict with normative rules.

The Technological Ecology

Alex Kirlik

University of Illinois at Urbana-Champaign, USA
kirlik@uiuc.edu

A variety of projects influenced by Brunswikian ideas have been completed in the past year. With my colleague Shin Maruyama (U. Tokyo), we now have an paper in press showing how probabilistic functionalism can be used to understand issues such as perception, coordination and improvisation in both music and human-technology interaction (*Proceedings of the IEEE*). The purpose of this work is to use group musical performance as an inspiration and model for the design of more robust and adaptive sociotechnical systems. We have also explored applications in education and intelligence. In education, we have co-authored (with R.J. Sternberg, D. Briney, L. Jarvin, S. Stemler & E. Grigorenko) a chapter on "Scaling up educational interventions," which uses representative design as a framework for understanding why various interventions do and do not result practically-relevant educational outcomes (to appear in R.J. Sternberg & M. Constas, Eds., *Translating Educational Theory and Research into Practice*, Erlbaum). In intelligence, we have written "Work in progress: reinventing intelligence for an invented world" (to appear in R. J. Sternberg & D. Preiss, Eds., *Intelligence and Technology*, Erlbaum). This work promotes a reconceptualization of human intelligence away from passive adaptation to "the environment" and toward a more dynamic and interactive perspective where both cognition and the environment are recognized as equally important objects of differentiation and modeling, and draws upon both Brunswikian and Gibsonian resources for ecological psychology. We have also recently had a theoretical piece "On Stoffregen's definition of affordances" accepted in the journal *Ecological Psychology*, with the premise that there is no magic bullet other than very detailed functional analysis for understanding the environmental properties participating in cognition and behavior. We are also pleased to have obtained Peter Pirolli's (PARC) book, *Information Foraging* as the second in the new

Oxford University Press series on "Human-Technology Interaction." (The first, largely Brunswikian-inspired, *Working with Technology in Mind*, is slated to appear in 2004). Pirolli's work (joint with Stuart Card) provides a theory of World-Wide-Web navigation based on animal models of foraging, John Anderson's rational analysis, and Brunswik's probabilistic functionalism. We also have three Brunswikian-oriented papers currently in review, on modeling airline taxi navigation errors (collaboration with Mike Byrne), and situation awareness in technological systems (collaboration with Richard Strauss), also available as Technical Reports from: http://www.aviation.uiuc.edu/UnitsHFD/report_fulltext.html

More about Confidence

Joshua Klayman

University of Chicago, USA
joshk@uchicago.edu

I am continuing to investigate the processes behind confidence, with several different sets of collaborators. Jack Soll (INSEAD) and I have recently completed studies showing that subjective confidence intervals (e.g., "I'm 80% sure that Tony Blair is between 42 and 65 years of age") are too narrow given the accuracy of people's knowledge, even when one samples items randomly from a domain and takes into account the possible biasing effects of unsystematic imperfections in judgment. The result is lots of overconfidence--although the format of asking for intervals has a large effect on how much. This contrasts with the two-choice case ("I'm 80% sure that Tony Blair is younger than the Dalai Lama"), where unrepresentative stimuli and unsystematic error seem to account for most of the apparent overconfidence bias. Jack and I are also conducting some studies to learn to what extent overconfidence is tied specifically to evaluation of knowledge and beliefs, vs. more general difficulty in understanding probabilities and distributions. In a related vein, we are also working with Peter Juslin and Anders Winman on a discussion of the role of limited and biased sampling of evidence in explaining overconfidence.

Katherine Burson (U. of Michigan) and Richard Larrick (Duke U.) and I have been studying a different kind of confidence, namely judgments about where one's performance or ability stands in relation to that of one's peers. Justin Kruger and David Dunning proposed a clever theory about this has recently gained a lot of publicity: People who don't know how to do something also don't know how bad they are, whereas good performers have a good sense of where they stand. This leads to general upward bias.

However, our studies show that inaccuracy in knowing where you stand relative to others is pretty well distributed across the board, and is not the domain of poor performers. In fact, the dominant effect is one also first noted by Kruger, namely that people tend to overestimate their standing only on tasks they find easy. On hard tasks, most people underestimate how their performance compares to others. This has been attributed to anchoring on one's own experience with insufficient adjustment for knowledge of others. However, Burson and I are currently conducting studies to investigate whether this can be better understood as simply a byproduct of people using perceived difficulty as a cue to their standing (following up on the work we talked about at last year's JDM conference.)

Between-Doctor Variation in COPD Gatekeeping Decisions

Olga Kostopoulou

Dept. of Primary Care & General Practice
University of Birmingham, UK
O.Kostopoulou@bham.ac.uk

Variation in clinical decisions is not an uncommon finding in research but its sources are difficult to pinpoint. We looked at how senior doctors decide to admit to intensive care or not a COPD* patient in extremis requiring mechanical ventilation. If respiratory support is not provided, some of these patients will die. However, if the disease has reached its terminal stage, ventilation will merely prolong the process of dying. Doctors therefore need to identify those patients who can receive 'sustained benefit in terms of quality and length of life'. This is a difficult task and there are not strong prognostic indices, variation was therefore expected.

Fourteen senior doctors (7 respiratory consultants and 7 intensive care consultants) from seven hospitals in the West Midlands participated. Six case histories were selected from hospital records of COPD patients. We employed a process-tracing approach, which consisted of

1. withholding case information until specifically requested by the doctors,
2. requesting survival estimates during the decision making process and
3. concurrent questioning of the doctors about interpretation of information and its impact on survival estimates and decisions.

The doctors' commentary was recorded and a written record was kept of what information was gathered in what sequence and at what point estimates or decisions were updated or finalised.

Most doctors agreed on the management of three of the cases: a relatively young patient with mild COPD (admit), an old patient with mild COPD (admit), and a patient with one functional limitation – needed help washing (do not admit). Decision variability concentrated on the other three cases that combined various good and bad prognostic factors, making them less clear-cut. Variability was attributed to doctors

1. requesting different information about a case,
2. weighting information differently and
3. interpreting information differently.

The numbers of respondents and cases were small, so our statistical results should be treated as preliminary. Admission decisions were associated with higher survival estimates but this was significant only in two patient cases. This is in agreement with other research that suggests that the link between prognostic estimates and clinical decisions may be less strong than is usually assumed.

There were significant differences between doctors in the amount of information gathered – which related neither to specific patient cases nor to the type of decision (admit/do not admit) nor doctor type (respiratory or intensive care). We do not know to what extent individual search strategies observed in the study reflect real-life behaviour or differing perceptions of urgency for 'paper patients'.

We argue for acknowledging and addressing all three potential sources of decision variation: identifying and interpreting correctly the clinical information relevant to the decision and assessing its relative importance. This necessitates identifying predictive indices and providing them to clinicians in the form of usable decision aids.

* COPD: Chronic Obstructive Pulmonary Disease. A progressive and irreversible respiratory disease with death rates similar to that of lung cancer

Sensor Noise and Ecological Interface Design: Patient Prioritisation for Elective General Surgery

Andrew MacCormick

Faculty of Medical and Health Sciences,
University of Auckland, New Zealand
andrew.maccormick@auckland.ac.nz

Our previous work involved performing judgment analysis on surgeons' judgments of patient priority for surgery using vignettes. In that work cluster analysis identified two philosophical groupings. Furthermore the poor agreement between surgeons in prioritising the vignettes was determined to be due to the poor consensus in weights applied to the cues.

We therefore undertook a feedback study to investigate whether consensus and therefore agreement could be improved. This was performed using a second round of the surgeons assessing the same vignettes with Cognitive Feedback in the form of the two different groups cue weightings.

A second project was the cross-validation of a cue weightings derived from nomothetic analysis of the initial data. This was to compare the use of vignettes and real patients in developing rule-based algorithms for the prioritisation of patients.

The initial analysis of the feedback study indicates that there was convergence towards consensus in the weighting of cues. The cross-validation indicates that there is poor agreement in the weighting given to cues when comparing vignettes and real patients. This is not surprising given the arguments regarding representative design and nomothetic versus idiographic analysis.

The use of Cognitive Feedback with surgeons assessing real patients is now being explored to develop a consensus approach to prioritisation and thus achieving patient equity of access to rationed health services.

Extinction and Habituation in the Context of Brunswikian Learning Theory

Erin C. McKiernan

ARL Division of Neurobiology,
University of Arizona, USA
emck@email.arizona.edu

Pedro Wolf

Dept. of Psychology, Ethology & Evolutionary
Psychology, University of Arizona, USA

Aurelio Jose Figueredo

Dept. of Psychology, Ethology & Evolutionary
Psychology University of Arizona, USA

Contemporary learning theorists have proposed that instinct and learning lie on opposite ends of a unidimensional continuum, requiring that the two be conceptualized as involved in a necessary tradeoff in which particular behaviors that concern a high degree of plasticity must inherently involve little preparedness, and vice versa (e.g. Rescorla & Wagner, 1972). Instead, Figueredo (1992, 1995, 2000) proposed a Brunswikian model in which the degree of learning and instinct involved in a particular behavior is the result of two independently varying parameters: (1) the average ecological validity of particular cues and means in the environment and (2) the variance around that average, over evolutionary time. Due to the emphasis on an evolutionary framework, one of the primary implications of this

theory is that the currently accepted dichotomy of associative versus nonassociative learning may not be a valuable construct (Figueredo, 1992, 1995, 2000). From this perspective, all learning is inherently associative. This leads to the prediction that behaviors such as extinction and habituation, commonly thought of as associative and nonassociative respectively, may be operationally and perhaps mechanistically identical. In other words, habituation can be thought of as extinction of a prepared association (Figueredo, 1992, 1995, 2000). The purpose of the current experiment is to provide evidence for the similarities between habituation and extinction by first examining the overt behavioral characteristics of both. We propose doing this in one comprehensive study of one model organism to supplement literature reviews, which have combined evidence from several studies and pointed out some similarities in need of more rigorous investigation (for a review see McSweeney & Swindell, 2002).

The model organism selected for this study was the butterfly *Pieris rapae*, in part due to its proboscis extension reflex (PER). This behavior, as shown in studies of this organism and others that also exhibit PER, can be both habituated and trained to allow for extinction (e.g. Braun & Bicker, 1992; Bitterman, 1983; Omura, Honda, & Hayashi, 2000). Butterflies will be exposed to several pairings of odor and sucrose to elicit a conditioned PER. Subsequent to conditioning, the animals will be randomly divided into two groups where one will undergo habituation of the unconditioned stimulus and the other will be extinguished to the conditioned stimulus. This way we can directly compare observable characteristics such as rate of responding, spontaneous recovery, contextual effects, etc. to elucidate whether the two learning processes can be viewed as operationally indistinguishable. Depending on the outcome of this study, future directions could include multiple experiments using various neurobiological methods to assess whether habituation and extinction use the same chemical pathways and neural substrates to generate learning.

Bitterman, M.E., Menzel, R., Fietz, A., & Schafer, S. (1983). Classical conditioning of proboscis extension in the honeybees (*Apis mellifera*). *Journal of Comparative Physiology* 97, 107-119.

Braun, G., Bicker, G. (1992). Habituation of an appetitive reflex in the honeybee. *Journal of Neurophysiology* 67 (3), 588-598.

Figueredo, A.J. (1992). Preparedness and plasticity: A stochastic optimality theory. Paper. Western Psychological Association Convention, Portland, Oregon, and the Young Investigator's Award Competition, Human Behavior and Evolution Society Meeting, Albuquerque, New Mexico.

- Figueredo, A.J. (1995). A stochastic optimality model of preparedness and plasticity. *Behavioral and Brain Sciences*, 18(2), 300-301.
- Figueredo, A.J. (2000). A Brunswikian optimality model for the evolutionary psychology of preparedness and plasticity. Paper. Brunswik 2000, Max Planck Institut für Bildungsforschung, Berlin, Germany.
- McSweeney, F.K., Swindell, S. (2002). Common processes may contribute to extinction and habituation. *The Journal of General Psychology* 129 (4), 364-400.
- Omura, H., Honda K., Hayashi, N. (2000). Floral scent of *Osmanthus fragrans* discourages foraging behavior of cabbage butterfly *Pieris rapae*. *Journal of Chemical Ecology* 26 (3), 655-666.
- Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory*, ed. AH Black, WF Prokasy, pp. 64-99. New York: Appleton-Century-Crofts

Brunswikian Approaches to Policy Analysis

Jeryl L. Mumpower

University Administration Building 417
University at Albany, SUNY, USA
mumpower@albany.edu

These days I spend most of my time making arbitrary and capricious decisions as a University vice president and dean. Occasionally I try to steal some time to do honest work and study decisions.

Tom Stewart and I have just completed a draft of a brief paper entitled "Asking "What If" Questions about Detection and Selection Processes in Public Policy -- An Application to Screening Mammography." This paper grew out of a session that I organized last year at the annual Association for Public Policy Analysis and Management meeting. The paper applies the Taylor-Russell model to address three critical policy questions:

1. What if the base rate for those who seek screening mammography were changed? The base rate might be changed in various ways, including recommending mammography only for older women, recommending mammography less frequently, and recommending mammography only for women with a family history of breast cancer.
2. What if the accuracy of screening mammography were increased? Ways of increasing the accuracy of mammography include improved reading procedures and improved image quality (e.g., digital mammography).

3. What if the decision threshold/selection rate were changed? This is under the control of the radiologist, yet has important policy implications that are rarely discussed.

I hope to continue to be able to squeeze out a little time to aid and abet my colleagues, especially Tom Stewart, in trying to bring Brunswikian approaches to policy analysis.

Recognition, rolling regression, control and insight

Ben Newell

University College London, UK
b.newell@ucl.ac.uk

Together with David Shanks and Tim Rakow we are continuing our empirical investigations of the use of "fast and frugal" heuristics. Recent experiments have included a follow-up on the examination of search strategies (reported in last year's newsletter) in which we have found more evidence that people are sensitive to both the ecological validity and the discrimination rate of cues in binary choice situations. We have also been taking a closer look at the "recognition heuristic" and testing its applicability in multiple-cue judgment tasks. The evidence so far seems to suggest that in multiple-cue tasks, recognition information is treated in the same way as other cues in the environment, rather than being attributed any 'special' non-compensatory status – contrary to the conception of recognition in the 'adaptive toolbox' framework. Both the search and recognition studies are in various stages of preparation and revision and we hope will appear in print before too long.

In work with David Lagnado and David Shanks we are investigating strategy development in multiple-cue tasks using a 'rolling regression' technique recently reported by Friedman and colleagues. The advantage of rolling regressions is that they capture the fluctuating state of the experimental environment as well as the evolution of individual's judgment strategies – aspects that tend to be lost when performance is aggregated across trials and compared with the regression weights characteristic of a complete training set. Initial results seem to indicate that in a four binary cue, binary outcome task people's judgment weights seem to track the environmental weights rather well. This fine-grained analysis of people's learning will, we hope, be able to demonstrate that the apparent use of simple strategies (or heuristics) and shifts between strategies, are emergent features of a single underlying mechanism which tracks the statistical properties of the environment.

Together with David Lagnado we are examining the role of intervention in multiple-cue learning. We are contrasting conditions in which participants are allowed to *intervene* and set the values of binary cues with standard conditions in which they *observe* pre-set values. In a simple task participants are required to learn whether a 'drawing program' will draw a circle or square based on the configuration of four buttons that can be set to either 'right' or 'left' – in the intervene condition the buttons are set by the participants, in the observe they are pre-set. In the judgment phase both groups rate the probability of an outcome conditional on a variety of test patterns. Initial results indicate an advantage for the interveners but only for configurations that were highly predictive of the outcome. We will present a poster based on this work at the meeting of the Society for Judgment and Decision Making in Vancouver.

Finally, Ben Newell is currently working with Peter Juslin in Uppsala on a project re-examining insight into judgment (see Karlsson & Juslin's comments in this newsletter). One aim of the project is to use the framework recently proposed by Juslin and colleagues to test predictions about the nature of insight. The framework suggests that shifts between performance based on mental cue-abstraction and integration, and exemplar memory can be induced by changing various aspects of the task environment. The idea is that such shifts in performance may also be accompanied by differences in the nature of insight into abstract and concrete aspects of both the task and the cognitive process – specifically that abstract knowledge should be better in conditions where judgments are based on cue-abstraction, but that concrete knowledge should be better when judgments are based on exemplar memory.

Shoe Print Investigation: The Process and Factors of Influence

Roos Paashuis
University of Amsterdam
paashuis@tm.tno.nl

José Kerstholt
TNO Human Factors
kerstholt@tm.tno.nl

Forensic researchers are more often required to justify their judgement in court. In my Ph.D. project I am examining the judgement process of Dutch forensic shoe print investigators, in order to assess the procedure they use, the cues they identify, the value they ascribe to these cues and the way they compare shoe prints to shoe tracks. Furthermore, I want to identify which factors affect these various

phases in the judgement process. A pilot experiment with one expert and two novices showed that several phases could be distinguished, such as an identification phase, an interpretation phase and an evaluation phase. These phases also revealed differences between the expert and novices. The expert, for example, identified cues with a higher characteristic value than the novices.

We are currently conducting a study in which a larger group of experienced and inexperienced investigators is compared. These investigators have to judge 4 real shoe print cases, two relatively simple and two relatively complex ones. The difference between a simple and a complex case is determined by the ease of abstracting cues from the track. In complex cases the track can be somewhat vague because of how the track was placed or secured. During task performance investigators have to indicate what the unique cues are, what value these cues have and how similar the cue is to the scratch in the shoe track. Eventually, insight into the judgement process of shoe print investigators can be used to secure knowledge for the organisation and to improve the training of shoe print investigators.

Nurses' and Psychiatrists' Judgements about Suicide Risk and Decisions on Observation: A Pilot Study

Brodie Paterson
University of Stirling, UK
b.a.paterson@stir.ac.uk

Dawn Dowding
University of York, UK
dawn.dowding@hyms.ac.uk

Clare Cassells
University of Stirling, UK
clare.cassells@stir.ac.uk

We have recently completed a pilot study designed to explore how psychiatrists and psychiatric nurses make judgements about suicide risk and decisions about observation levels in acute inpatient settings in the UK. The study utilizes social judgement analysis for this purpose.

4 psychiatrists and 17 psychiatric nurses from one Scottish health care trust comprised the sample for this study. Each participant was given a vignette booklet consisting of 145 vignettes (including 15 repeat cases for consistency analyses) to complete. Information presented in the vignettes was varied for 13 cues of interest: Suicidal ideation, previous suicide attempts diagnosis, comorbid substance abuse, length of admission, hopelessness, insight, sex,

protective factors, adverse life events, previous no. of admissions, patient compliance and clinical improvement. For each vignette, participants were required to judge risk of suicide occurring within 24 hours on a visual analogue scale (ranging from very low risk to very high risk). Subsequently, they were asked to recommend 1 observation category from a list of 4 (pass, general observation, constant observation and special observation) for each vignette. Stepwise linear regressions were carried out on the data to obtain professional group judgement policies and individual judgement policies.

The psychiatrist group analysis produced a model with 7 significant predictor variables, explaining 30% of the variance in risk judgements ($F_{7,510} = 32.60$, $p < 0.001$). Significant variables in order of importance (determined from beta weightings) were suicidal ideation, diagnosis, previous suicide attempts, patient sex, hopelessness, clinical improvement and length of admission. In comparison, the nurse model analysis produced a model explaining 18.2% of variance with 7 significant predictor variables ($F_{7,2209} = 70.8$, $p < 0.001$). Significant variables were suicidal ideation, previous suicide attempts, compliance, length of admission, clinical improvement, patient sex and hopelessness. The relative beta weightings in both groups indicate that suicidal ideation and suicide attempts are strongly related to suicide risk judgements. In terms of differences, diagnosis is considered an important predictor of risk in the psychiatrist group, but not in the nurses group. Nurses instead would appear to attach more importance to length of admission. In both instances, length of admission had a negative weighting, indicating that perceived risk of suicide decreases with lengthier admissions. These differences in cue importance also emerged at an individual level analysis.

Decisions on observation were comparatively similar for both groups. Proportion of patients placed on each observation level for psychiatrists and nurses respectively are as follows: Pass (4.1%, 3.8%) General Observation (63.3%, 73.4%) Constant Observation (22.2%, 22.2%) Special Observation (10.4%, 0.5%). There is a slight tendency for psychiatrists to recommend the use of special observation more frequently than nurses. Stepwise linear regressions analyses revealed that risk judgements had a large influence on observation decisions. The variance explained by perceived risk judgements in decision-making ranged from 49% to 83% for psychiatrists and from 11% to 81% for nurses.

Consistency in judgements and decisions was also assessed for each individual using related samples tests. Both psychiatrists and nurses showed a 75%

consistency level for risk judgements for identical cases presented at Time 1 and Time 2. (In this instance, 3 t-tests for psychiatrists and 13 t-tests for nurses failed to approach statistical significance, highest $t = -1.516$ smallest p -value = .152). The results for individuals who were inconsistent (1 psychiatrist and 4 nurses) were analyzed further to determine whether observation decisions also differed at time 2. Although all participants ($n=5$) significantly reduced their risk perceptions at time 2, only 1 participant (1 psychiatrist) reduced observation levels at time 2 (from higher levels to lower levels (Wilcoxon = -2.121, $p = 0.034$)) For the remaining participants, observation levels remained largely unchanged.

Overall, the results from this pilot study give us some insight into how judgements and decisions about suicide risk are made in acute inpatient settings. Degrees of similarity and variation exist between professional groups and individuals in terms of what factors are attended to when making risk judgements and decisions. One major commonality is a tendency to rely on standard predictors of risk which (according to the literature) may not be effective short term predictors of risk. This has particular clinical implications since the concept of suicide occurring in the short term is integral to practice. Overall consistency levels are high, which means that patients on the ward with clinically similar presentations will be judged to be at similar risk and treated in similar ways.

At present we are undertaking the main phase of the study, which is being extended to other health care trusts in Scotland. This will allow us to make more comparisons and ultimately compliment the findings we currently have.

Expert Witness Use of Consistency Cues: Judging the Veracity of Testimonies

Callia Piperides

Max Planck Institute, Berlin, Germany
piperides@mpib-berlin.mpg.de

A scientifically validated method used to assess witness credibility is *content analysis*, which assesses the semantic content of a written testimony. Consistency analysis, a component of this analysis, involves a comparison of witness statements made at successive interviews regarding contextual discrepancies (such as, consistencies, contradictions, omissions, and additions) in the events described. In Germany, content analysis can be a powerful piece of evidence in courts, which can affect the outcome of the case if no other incriminating or exonerating evidence is available. In the US, the method (known as *statement validity assessment*) is used in assisting

with police investigations and as an educational tool in the courts to aid in understanding the quality of evidence. The consequences of wrongful classifications can have devastating effects for both victims and defendants.

The objective of the study was to investigate how accurate psychologists, acting as expert witnesses, are in classifying statements. In addition, as expert witnesses have considerable discretion, their decision-making was investigated to better understand how they use consistency cues. Four experienced psychologists in content analysis analysed 22 statement sequences (44 statements) made by 11 German psychology students. At a first interview, the student participants made two lengthy and detailed statements each: a false narrative based on fantasy and a true one based on an experienced event. These statements were repeated a year later. All interviews were recorded and transcribed.

The analysis of cue usage showed that experts rated five cues of eleven as important to their credibility judgements. Of these cues, "contradictions in the core events" of a narrative was rated as the most important cue for both positive and negative credibility judgements. A lens model analysis (Brunswik, 1952) supported these findings and showed that a simple model with only one significant discriminating cue—contradictions in the core events—best described both the expert judgements and the veracity of statements. Experts' overall accuracy rate in classifying statements was 66%. More specifically, the findings showed that whereas true statements were classified very well, false statements were underestimated and correctly classified at only above the chance level.

The findings of the study showed that a consistency analysis can improve performance in the classification of false statements and can reduce uncertainty in judgements. In addition, it was shown that "contradictions in the core events of narratives" as used by experts is a valid credibility cue in distinguishing between true and false statements.

Conclusions: Seen in practical terms, this accuracy level is based on only a third of the information normally made available to experts assessing statement credibility and should be evaluated in this light. The consequences of being better at classifying true statements and underestimating false cases will depend on whose testimony is being evaluated. The findings of the present study are important in that they show that a consistency analysis can improve performance in the classification of false statements and reduce uncertainty in judgements. In addition, it was shown that "inconsistencies in the core events of narratives" as used by experts is a valid credibility cue in distinguishing between true and false

statements. Future research needs to be directed at investigating the validity of other consistency cues in order to improve expert performance.

Latent Problem Solving Analysis (LPSA): A computational theory of representation in complex, dynamic problem solving tasks

Jose Quesada

Institute of Cognitive Science
University of Colorado (Boulder), USA
quesadaj@psych.colorado.edu

Many real-world decision making and problem solving situations are (1) dynamic, because early actions determine the environment in which subsequent decision must be made, and features of the task environment may change independently of the solver's actions; (2) time-dependent, because decisions must be made at the correct moment in relation to environmental demands; and (3) complex, in the sense that most variables are not related to each other in a one-to-one manner. However, traditional, experimental problem solving research has focused largely on tasks such as anagrams, concept identification, puzzles, etc. that are not representative of the features described above. Several researchers have started working on a set of computer-based, experimental tasks that are dynamic, time-dependent, and complex, called microworlds, and the area of thinking and reasoning that deals with them has been called Complex Problem Solving (CPS). The richness of the data generated when participants are solving microworld problems is simply overwhelming. For years, CPS has been plagued with methodological problems, and performance has been analyzed at a very shallow level (process measures were mostly discarded). The interest of this research paradigm, a hybrid between field studies and experimental ones, is tied to the success of methodological advances that enable performance to be analyzed. This work introduces a new, abstract conceptualization of microworld research using an approach based on Latent Semantic Analysis (LSA), called Latent Problem Solving Analysis (LPSA).

LSA is a machine-learning model that induces representations of the meaning of words by analyzing the relation between words and passages in large bodies of representative text. LSA is both a method (tool) used to develop technology to improve educational applications, and a theory of knowledge representation used to model well known experimental effects in text comprehension and priming, among others (Landauer & Dumais, 1997). Latent Semantic Analysis was originally developed in the context of information retrieval (Deerwester,

Dumais, Furnas, Landauer, & Harshman, 1991) as a way of overcoming problems with polysemy and synonymy. Some words appear in the same contexts (synonyms) and an important part of word usage patterns is blurred by accidental and inessential information. The method used by LSA to capture the essential semantic information is dimension reduction, selecting the most important dimensions from a co-occurrence matrix decomposed using Singular Value Decomposition. As a result, LSA offers a way of assessing semantic similarity between any two samples of text in an automatic, unsupervised way.

Although LSA has been mostly used on text corpora, LSA's basic approach can be applied to any domain of knowledge where there are a high number of weak relations between tokens, as in Complex Problem Solving (tokens representing states of the system or actions to control it). Instead of using word co-occurrence statistics and huge samples of text, we have used a representative amount of activity in controlling dynamic systems, and actions or states have been used to develop the much-wanted objective measure of similarity for changing, time-dependent, highly complex experimental tasks. Like words, states and actions appear in particular contexts but not in others. Some states and actions are interchangeable, being "functional synonyms" Given the right algorithms and sufficient amounts of logged trials, a problem space can be derived in a similar way as semantic spaces are. This analysis (LPSA) proposes that the problem spaces that people use to represent their environments are derived empirically from experience, and can be approximated using a vector space formalism. The implications for current theories of problem solving and representation are discussed.

This work has four main purposes: (1) Solve the methodological problems associated with microworld performance assessment to enable rigorous experimental manipulations in highly complex laboratory tasks. (2) Propose an alternative theory of problem solving and representation in these domains that is mostly bottom-up, and compare it to existing proposals. (3) Present our approach as a computational theory of expertise that is able to reconcile results that were previously explained only by substantially different theories. (4) Develop a landing technique automatic assessment system based on our theoretical framework, using a high fidelity flying simulator, to show how our approach can be applied to real-world problems. To prove the theoretical power of the LPSA framework, we applied it to very different complex tasks: It is a good source of evidence for the psychological reality of a theory if it can explain very different domains without having to change its assumptions. The main conclusions and

results are thus presented for each of the different complex tasks that we used:

(1) Firechief is a spatial microworld where participants control helicopters and trucks to extinguish a fire that spreads in a forest. Previous literature (Cañas, Quesada, Antoli, & Fajardo, in press; Quesada, Cañas, & Antoli, 2000) showed that people have trouble adapting their problem solving activity to new environmental situations; LPSA applied to those experiments adds new insights about what make people stick to old solutions that are not efficient when the environment changes. Comparisons with human judgments validate LPSA similarities very well: when humans are asked to say how similar two trial videos are, their opinion correlates .94 with the LPSA cosines for these trials.

(2) Duress is a thermodynamic microworld that simulates hydraulic process control, designed to be representative of industrial process control systems. It consists of two redundant feedwater streams that can be configured to supply either, both or neither of two reservoirs. We used the data from a six-month long, six-participants experiment reported in, Christoffersen, Hunter, and Vicente, (1996, 1997, 1998) to generate a "simulated expert" with three years of experience with the system. Our system could predict the last quarter of any trial using the first three quarters with an average accuracy of .8. When the system was given an experience of only six months the predictions fell down to less than .3. If the system is trained with three years of practice in an environment with no constraints (that is, not governed by rules of conservation of mass and energy), the predictions were not useful as expected, and comparable to the novice level. LPSA's explanation for these results was able to generate predictions that either process theories or product theories of expertise could explain; both have never been explained by a single computational theory.

(3) High-Fidelity Flying simulator, used to develop a landing technique automatic assessment system based on LPSA. We collected 400 landings where the landing conditions were manipulated systematically, and created a vector space with them. Two instructors evaluated the landing, one of them sitting in the copilot seat, and the other one watching plots of relevant variables in real time (complete and reduced information experts respectively). The model was trained with the variables that the reduced information expert used in his plots to evaluate the landing technique. Then, the nearest neighbors of any new landing were used to generate automatic ratings. The ratings that the model emitted agreed with both humans as much as the two human graders agreed with each other.

We present a theory of representation during experienced human problem solving, centered in the framework and results described above, that has a strong bottom-up component. Most computational theories of problem solving are based on a symbolic rule-based approach, where the relationship between states is defined explicitly (e.g., IS-A, PART-OF, SAME-LIST). The learning algorithms proposed are oftentimes centered on chunking mechanisms. In our approach, the elementary units are located in a metric space, and their combinations are based on arithmetic operations. Although partial, (we propose it only for problem solving tasks that contain a huge vocabulary and have been learned during a long time), our explanation of problem solving as navigation in a metric space where functionality is the main grouping criterion is radically new.

Cañas, J. J., Quesada, J. F., Antoli, A., & Fajardo, I. (in press). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem solving tasks. *Ergonomics*.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1991). Indexing By Latent Semantic Analysis. *Journal of the American Society For Information Science*, 41, 391-407.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.

Quesada, J. F., Cañas, J. J., & Antoli, A. (2000). An explanation of human errors based on environmental changes and problem solving strategies. In C. P. Warren (Ed.), *ECCE-10: Confronting Reality*. Sweden: EACE.

The Use of Multi-objective Optimization toward Development of a Noncompensatory Lens Model

Ling Rothrock

The Pennsylvania State University, USA
lrothroc@psu.edu

To account for human decision making in time-stressed, information-rich tasks, Rothrock and Kirlik (2003a) present a technique, called Genetics-Based Policy Capturing (GBPC), for inferring noncompensatory, rule-based heuristics from judgment data, as an alternative to regression. In GBPC, rule-base representation and search uses a genetic algorithm, and fitting the model to data uses multi-objective optimization to maximize fit on three dimensions: a) *completeness* (all human judgments are represented); b) *specificity* (maximal concreteness); and c) *parsimony* (no unnecessary rules are used). GBPC is illustrated using data from

the highest and lowest scoring participants in a simulated dynamic, combat information center (CIC) task. GBPC inferred rule-bases for these two performers that shed light on both skill and error. The GBPC results were compared with regression-based Lens Modeling of the same data set, and enabled the authors to interpret the high scoring performer's highly significant use of unmodeled knowledge (C=1) revealed by Lens Model analysis. The GBPC findings also allow the authors interpret a similarly high use of unmodeled knowledge (C=1) in a previously published Lens Model analysis of a different data set collected in the same experimental task (Bisantz et al., 2000).

GBPC provides the basis from which an integrated model of human-environment system, called the Genetics-based Lens Model or GBLM, can be formed. An initial study to investigate the utility of the GBLM (Rothrock and Kirlik, 2003b) used the framework shown in Figure 1 to compare GBLM and regression model performance under different combinations of environment cue structures and judgment strategies. In Cases I and IV, the organizing principle of the cues is linear. In contrast, Cases II and III represent nonlinear cue-criterion relationships and are characterized by rule-based descriptions. The analog to the cue-criterion distinctions exists for the judge in terms of compensatory and noncompensatory decision strategies. A compensatory strategy, such as Cases I and III, represents a cue-judge relationship that can be characterized by a linear formulation. In contrast, a noncompensatory strategy, such as Cases II and IV, is exemplified by conjunctive and disjunctive rules.

Cue Structure

	Linear	Nonlinear
Decision Strategy		
Compensatory	Case I	Case III
Noncompensatory	Case IV	Case II

Figure 1. Framework to Investigate Decision Making Models

The present study focused specifically on Cases I and II. First, Case I is the ideal case for the use of the Lens Model Equation (LME). If both the judgment and criterion models are linear, it is expected that the amount of unmodeled knowledge, or C, is minimized. On the contrary, Case II presents the worst case for using the LME. That is, the linear knowledge accounted by the model is expected to be minimal

and C is maximized. In fact, the results of the study showed that GBLM was able to account for cognitive control (R_s) and environmental predictability (R_e) in Case II better than its linear counterpart. However, “unmodeled” knowledge remained high in the Case II GBLM analysis. This phenomenon offers opportunity for future research.

The GBPC problem in Rothrock and Kirlik (2003a) was formalized into a 0-1 multi-objective linear programming (MOLP) problem by Rothrock, Ventura and Park (2003). The objectives of the formalization were to extend the solution method beyond genetic algorithms and to understand the complexity of the problem. The researchers provided sensitivity analyses of exhaustive solutions to simple decision problems and propose the use of heuristic algorithms for more complex problems. For example, in the case of a noncompensatory strategy consisting of four rules of length four, the objective function must be computed 1,663,740 times. For a Pentium-IV class PC, this consumes over seven hours of processing time. The example is relatively small in size compared to the encoding for a complex decision making task. For the command and control problem discussed in Rothrock and Kirlik (2003a), an exhaustive search would require the objective fitness function to be computed 2.107×10^{87} times. The exhaustive search technique, therefore, is prohibitively and alternative heuristic algorithms in addition to genetic algorithms are currently under investigation.

Bisantz, A. M., Kirlik, A., Gay, P., Phipps, D., Walker, N., & Fisk, A. D. (2000). Modeling and Analysis of a Dynamic Judgment Task Using a Lens Model Approach. *IEEE Transaction on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 30(6), 605-616.

Rothrock, L., & Kirlik, A. (2003a). Inferring rule-based strategies in dynamic judgment tasks: toward a noncompensatory formulation of the lens model. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 33(1), 58-72.

Rothrock, L., & Kirlik, A. (2003b). *Explorations in Modeling Human Decision Making in Dynamic Contexts*. Paper presented at the 5th International Conference on Engineering Psychology and Cognitive Ergonomics, Crete, Greece.

Rothrock, L., Ventura, J., & Park, S. (2003). *An Optimization Methodology to Investigate Operator Impact on Quality of Service*. Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (IITSEC), Orlando, FL.

Metacognition in Multiple Cue Probability Tasks

Peter Schaefer

University of Georgia, USA
schaeferp1874@yahoo.com

In the past year I completed my Ph. D. under the direction of Rob Mahan at the University of Georgia. The dissertation was an investigation of insight and metacognitive ability within the context of Cognitive Continuum Theory. In two experiments, participants completed two standard measures of metacognition. They then completed an MCPL task in one of eight conditions. (I constructed environments that were either analytical or intuitive in nature. This manipulation was then crossed with four function forms [2 linear, 2 quadratic]).

The main queries of the investigation were 1) does insight, in fact, covary with task type as predicted by CCT? 2) does insight covary significantly with metacognition, measured independently (thus assessing the sometime equation of insight with metacognition)? 3) does metacognition predict performance more in some task types than others?

Zero-order across-condition correlations in both experiments indicated that performance (most notably cognitive control) was significantly correlated with the metacognitive measures. Interestingly, the correlations between metacognition and insight were strongest in the linear conditions in both experiments, and were largely attenuated in the quadratic conditions. An interrelated finding is that the roles of insight and metacognition as predictors of performance varied, again as a function of linear vs. quadratic function forms. In both experiments, insight alone predicted unique performance variance in the linear conditions whereas in the quadratic conditions metacognition also predicted unique performance variance. One potential interpretation of this is that when tasks are linear in nature insight is a task-specific measurement that captures the skills necessary for consistent application of a strategy. Metacognition, to the extent that it predicts performance, appears to be largely superfluous. When tasks are nonlinear in nature, insight is harder to come by (contrary to CCT, insight was significantly lower in the quadratic conditions). Therefore, the broader repertoire of behaviors included under the metacognitive umbrella (e.g., comprehension checking, strategy evaluation, etc.) may be more germane.

Possible extensions include a more thorough systematic manipulation (following the dreaded “law of one variable”) and the inclusion of more measures of metacognition. It might be of interest to see what measures or aspects of metacognition insight “loads” onto.

In a different vein, I am planning another potential project involving the use of the Big 5 model of personality in MCPL, specifically with regards to the link found between extraversion & overconfidence, and openness to experience/intellectance and better calibration (Schaefer, Williams, Goodie, & Campbell, in press).

Schaefer, P.S., Williams, C.C., Goodie, A.S., & Campbell, W.K. (in press). Overconfidence and the Big Five. *Journal of Research in Personality*.

The Statistical Structure of Speech Sounds Predicts Musical Universals

David A. Schwartz

Duke University, USA

Schwartz@neuro.duke.edu

The similarity of musical scales and consonance judgments across human populations has no generally accepted explanation. In a recently published paper (Schwartz et al., 2003) we presented evidence that these aspects of auditory perception arise from the statistical structure of naturally occurring periodic sound stimuli. An analysis of speech sounds, the principal source of periodic sound stimuli in the human auditory environment, shows that the probability distribution of amplitude/frequency combinations in human utterances predicts both the structure of the chromatic scale and consonance ordering. These observations suggest that what we hear is determined by the statistical relationship between auditory stimuli and their naturally occurring sources, rather than by the physical parameters of the stimulus per se. The work thus extends to audition and esthetics the methodological approach Brunswick pioneered a half-century ago in his work on contour grouping in vision.

Schwartz DA, Howe CQ, and Purves D (2002) The statistical structure of speech sounds predicts musical universals. *J Neurosci* 23(18):7160-7168.

Read a lay language summary of the research at <http://www.newscientist.com/news/news.jsp?id=ns99994031>.

Overconfidence vs Self-Confidence in Experts

James Shanteau

Kansas State University, USA

shanteau@ksu.edu

Denis Hilton

University of Toulouse, France

We are conducting a comparison of overconfidence in the judgments of experts vs self-confidence as a behavioral characteristic of experts. First, some background.

OVERCONFIDENCE is defined as "an unwarranted belief in one's correctness" (Lichtenstein, et al, 1977). As such, it reflects an "excessive narrowness of subjective probability distributions" (Anderson, et al, 1981). Confidence is often assessed by calibration, where "a judge is calibrated if, over the long run, the proportion that is true equals the probability assigned" (Lichtenstein, et al, 1982).

Overconfidence is routinely found in novices, eg, Alpert and Raiffa (1969) found that answers given .90 confidence were 75% correct and answers given 1.00 confidence were 80% correct. Moreover, overconfidence is often seen for moderate and difficult items, while underconfidence is seen for easy items -- the hard-easy effect.

Similar results have also been observed in experts. For example, Oskamp (1962) observed overconfidence in expert psychiatric diagnoses for hard cases, but underconfidence for easy cases. An exception, however, is that expert weather forecasters are well calibrated, eg, Murphy and Winkler (1974) found an average deviation of .028 from perfect calibration for temperature forecasts.

SELF-CONFIDENCE has been observed as a consistent behavioral characteristic for many types of experts. Shanteau (1987) proposed that experts share many psychological attributes, such as strong communication skills, ability to tolerate stress, high creativity, ability to simplify complexities, and high outward self-confidence in their decisions. When asked, experts routinely report strong self confidence (Shanteau & Abdolmohammadi, 1992). However, there are domain differences with higher self-confidence seen for doctors than for nurses or accountants (Shanteau & Peters, 1989).

Abdolmohammadi, et al (in press) extended previous research by asking expert auditors to assess the importance of various attributes of expert auditors. While knowledge and experience were the most highly ranked, self-confidence ("a strong belief in his/her ability to make good decisions") was also in the top group of characteristics (6 out of 25).

QUESTION: What is the connection, if any, between overconfidence in the judgments of experts and self-confidence as an attribute of experts? Put another way, is there a necessary connection between overconfidence and self-confidence? Or are these independent properties?

SOME ANSWERS: Self-confidence lets experts deal with the responsibility of making major decisions by "shaking off" the consequences of bad outcomes. Also, self-confidence may be an essential component to the public presentation of experts (Goffman, 1959).

At the same time, overconfidence may be part of the desire of experts to be seen as knowledgeable and decisive. After all, clients expect experts to make highly confident decisions -- why else would they hire them?

COMMENTS: First, unwarranted confidence is clearly undesirable, especially in decisions of experts. Money and, more importantly, lives may depend on these decisions. Having appropriate information about level of confidence (high or low) is therefore critical.

Second, an expert generally needs to have a strong personal belief in his/her own abilities. Otherwise, the expert would be unable to recover from negative outcomes, eg, the death of a patient. Also, experts who seem less than certain (eg, "I am 80% confident") are unlikely to have their advice followed by clients.

Third, in legal settings expert witnesses are told to avoid using the language of uncertainty (eg, "the data shows with 95% confidence that..."). Instead, they are told to sound as certain as possible (eg, "the data proves..."). Thus, a job requirement for legal experts is to overstate the confidence in their decisions.

ONGOING RESEARCH ISSUES: We are examining the relation between overconfidence and self-confidence in both previous and new studies of experts. We also plan to study the evolution of overconfidence and self-confidence as novices develop into experts. Finally, we will be comparing overconfidence and self-confidence across different domains of expertise, eg, natural science vs behavioral science.

FINAL ISSUES: When is overconfidence NOT a problem? If only the rank order of options is important to clients, then it does not matter if recommendations by experts are stated with greater than warranted certainty.

When is overconfidence a problem? If clients make use of the precise values stated by experts, then overconfidence can be misleading as to the certainty of the decisions.

Lastly, are self-confidence and overconfidence necessary properties of experts? If so, then it may not be possible to have one without the other?

Clinical Guidelines on Depression: A Qualitative Study of GPs' Views

Liz Smith

University of Aberdeen, UK
mes@hsru.abdn.ac.uk

Anne Walker

University of Aberdeen, UK

Ken Gilhooly

Brunel University, UK

This was a qualitative study, which I did to enhance findings from a previous SJT research. Clinical guidelines are a familiar component of health care, although passive dissemination of them does not ensure their implementation. This study examines General Practitioners' (GPs) views of guideline implementation, particularly about guidelines for the management of patients with depression. In-depth interviews were used with a purposive sample of 11 GPs from the Grampian region of Scotland, and the North East of England who had all participated in a previous SJT study. Data were analysed using the 'framework' technique.

The previous questionnaire, identified factors which influenced GPs prescribing decisions and how these differed from those the guidelines recommend. Brunswik's lens model was employed to explore individual treatment decision policies of GPs for 20 case vignettes. Individual policies were then aggregated and compared with those derived from guideline recommendations. Important differences emerged between the two in the utilisation of cues and there was considerable variation between GPs' decision policies and guideline compliance. Guidelines placed more importance on the duration of symptoms whereas GPs gave weight also to other symptoms. GPs prescribed antidepressants at a greater rate than was recommended by the guidelines.

The main priorities of this qualitative study are to address the questions that were raised:

- (1) Do GPs fail to follow this particular guideline because they do not agree with the recommendations made by the depression guideline?
- (2) Do they not follow it because they feel that guidelines are insufficiently flexible to

manage the variety of patients who have symptoms of depression?

- (3) What barriers do GPs think stop them from following guideline recommendations?
- (4) What would be the most fruitful method to promote increased guideline use?
- (5) Why was there so much variance between the GPs' compliance?

Two GPs had no particular problems with the recommendations. However some of the GPs did not agree with some part of the guideline, which may account for the previously reported variable compliance. One area of disagreement was with the recommendation to refer patients, as specialists were not available for this. Referring patients was a recurring issue and is discussed below. Another area of disagreement was the duration of symptoms criterion

“ The guidelines stipulate that you have to have these features and for at least two weeks ...and if they only had it for a week why should I wait for a week, why should they be miserable for a week, when I am pretty certain they are depressed?” (GP3).

Many of the interviewed GPs thought the guidelines were insufficiently flexible to use with the variety of depressed patients they see. Depressed patients present with varying pre-existing illness combinations, beliefs and attitudes about depression and its care, treatment preferences, concerns about confidentiality and stigma as well as varying degrees of access to care. One GP talking of his worry about lawyers becoming involved in checking guideline compliance, said that this would result in defensive practice rather than providing the best treatment for the individual patient. This was echoed by another GP who said that guidelines should not be used in all situations as they vary so much and one GP that thought the guidelines made invalid assumptions about patients' illnesses and GPs' time.

One of the commonest perceived barriers which prevents these GPs from following guidelines is the volume of guidelines they receive. They feel that they receive too many to have time to read them all and do not always know which one to follow. They gave the impression of feeling despondent and overwhelmed.

“...There's a bit of numbing in there as well, oh no, not another guideline.” (GP11)

“We get flooded with stuff ... With a lot of stuff I bin it or file it.” (GP5).

Linked to the volume of guidelines received was 'lack of time'. GPs consistently viewed this as a barrier to

guideline use. A related issue was that GPs thought guidelines were not always accessible. Lack of resources re-emerged as a major barrier to following guideline recommendations. Problems were reported about referring patients ranging from having no specialist to refer them to, the patients being misled about the specialists' qualifications, patient confidentiality issues and long waiting times. Several GPs reported that by the time the patient had the appointment the problem had gone and they no longer wanted it

“...a guideline might come through and I've followed the protocol...and arranged a referral...then the reply has come back from the hospital that they don't have the resources for this at the moment. So it [the guideline] has all fallen flat on its face and that is extremely disappointing when we in primary care are trying our best .” (GP2).

In order for guideline use to increase, GPs in this study would like to see: more resources put in place; a reduction in the number of guidelines they receive; guidelines being produced by and sent from one central body which would have a multidisciplinary team including some GPs; incorporation of guideline recommendations into computer decision support systems; and regular audit and feedback to allow them to monitor their practice. This study shows that a range of factors probably contribute to variability in GP compliance with guidelines for the management of depression. One obvious implication from this study is that before recommendations are made, such as to send certain patients to specialists, they need to make sure that these specialists are available. Or perhaps the role of guidelines is seen by those involved in their production as one to push for more mental health professionals. If we are serious about closing the gap between research evidence and practice, it may be that a strictly new system of guideline development is needed at a national level where there is just one national centre where a multi-disciplinary team, including some GPs, is responsible for the production of guidelines, dissemination of a reasonable number of guidelines, incorporation of them onto computer systems, auditing performance and giving feedback to GPs. This study has opened up possibilities that should be further explored.

Brunswik's Influence on Geographic Profiling Research

Brent Snook*

University of New Brunswick, Canada;
bsnook@unbsj.ca

Paul J. Taylor

University of Liverpool, UK
pjtaylor@liv.ac.uk

Craig Bennell

Carleton University, Canada
cbennell@connect.carleton.ca

We decided to remain consistent with the structure of articles in previous Brunswik Society Newsletters by providing a summary of our research on a police decision-making task known as *geographic profiling*. This task requires a person to make a prediction about the location of an offender's residence based on information about where that offender committed his or her crimes. While we have yet to directly investigate whether or not probabilistic regularities govern the mechanism people use to perform this task, our investigations have been heavily influenced by Brunswik's philosophy of psychological research. We explain below how Brunswik's thinking has influenced our approach to studying consequential real life problems.

Perhaps the largest influence that Brunswik has had on our research is to enforce the importance of designing experiments that carry meaning beyond the laboratory. In that vein, we have aimed to ensure that our research is of importance for both public safety and policy issues and that our experimental results are generalisable to future investigative settings. This has required that we embrace Brunswik's notion of representative design, since offenders' actions are inevitably the consequence of complex, random interactions with the environment. Specifically, in our experiments, we have exposed participants to stimuli that is randomly sampled from large databases of actual serial offender information rather than artificially designed or selected. In measuring responses, we have similarly tried to limit the extent we constrain participants responses, often asking for free-response explanations of how they are tackling the task.

While we did not explicitly use Brunswik's Lens Model as a framework for our manuscript, a second impact of his thinking has been to provide a useful structure for considering the link between the cognitive strategies used by our participants (i.e., organism) and the patterns of offender spatial behaviour (i.e., environment). For example, in one study we tested the functional achievement (i.e., accuracy) of 215 participants on a task that required

them to predict the residential locations of 10 randomly sampled serial offenders before and after being introduced to some cognitive strategies. Recognising the importance of introducing strategies that match the empirical regularities found in offender spatial behaviour, we introduced two strategies: the Decay strategy, which predicts that many serial offenders live close to their crime locations, and the Circle strategy, which predicts that many offenders live within a circle with its diameter defined by the distance between the offender's two furthestmost crime locations. The Decay strategy matches the long-established finding that offenders do not travel far from their home to offend and that the frequency of offending follows an inverted "J" distribution (i.e., there are many more targets selected closer to, rather than further from, an offender's home). The Circle strategy matches evidence showing that the majority (often over 80%) of violent serial offenders' homes are located within an area demarcated by their two most distant crimes.

Our aggregate-level analysis of this study showed that participants were significantly less accurate in their predictions compared to the prescribed actuarial technique. However, once participants were introduced to one of two decision-making strategies (Decay or Circle), there was a significant improvement in the accuracy of their predictions, to a point where post-training performance did not differ significantly from the predictions of the actuarial technique. Yet, this analysis is limited in its usefulness, since it gives us no clues as to whether or not all participants were able to utilise the strategies, nor whether or not these strategies were effective in every map. Thus, in true Brunswikian fashion, we turned to an idiographic analysis of performances. This analysis showed that half of our participants made accurate predictions before being provided with the strategies. By looking at the relationship between the cognitive strategies used by the participants (before training) and the regularities of the offenders spatial behaviour (e.g., how central the sampled offenders live in relation to where they select their targets), we found that higher levels of predictive accuracy was the result of a match between the cognition of participants and the structure of the decision environment. Such detailed findings are not commonplace to our literature.

Overall, Brunswik's emphasis on the necessity to conduct research that will be meaningful outside the laboratory, to consider the representativeness of ones experimental design, and to explore the relationship between an organism's behaviour and mental processes and the structure of the environment has proven invaluable in our research. We plan to continue exploiting Brunswik's ideas in future research looking at the geographic profiling

task (and other forensic related decision-making tasks) and to contribute to his notion of probabilistic functionalism.

**The order of authors is not meaningful.*

Effects of Data Uncertainty in a Process Control Microworld

Olivier St-Cyr and Kim J. Vicente

Cognitive Engineering Laboratory
Department of Mechanical & Industrial Engineering,
University of Toronto, Canada
olivier.st.cyr@utoronto.ca

The purpose of our Brunswik-related research is to investigate potential effects of sensor noise on the Ecological Interface Design (EID; Vicente and Rasmussen, 1992) framework. In particular, we want to understand how sensor noise will affect operators' control strategies and performance.

Modern process control plants incorporate several sensors that are installed at strategic locations throughout the system (Johnson, 1997). Despite the current state of technology in sensing devices, information transmitted by the instrumentation and control equipment is often noisy (Stein, 1969). Therefore, data about the state of world will be uncertain, potentially affecting both the display content and the ways operators will control the equipment.

When information about the world is inexact, operators may have to adjust their decision-making tactics to account for the uncertain data. Collecting and integrating data which reflect uncertainty (e.g. sensor noise) will require high cognitive demands. Moreover, since the data is in part unreliable, different strategies might have to be explored to control the system efficiently (Woods, 1988). Uncertainty can also affect decisions and actions made by operators since the data have the potential of losing their real meanings and more interpretation might be required (Finger & Bisantz, 2002). Based on these indications, there seems to be a connection between sensor noise and control strategies (i.e. the ability for operators to adapt to uncertainty in the environment).

To understand to effect of sensor noise on operators' control strategies, a series of studies will be conducted on the DUal REservoir Simulation System II (DURESS II). Two different interfaces (P and P+F) for the same microworld were developed (Vicente & Rasmussen, 1990; Pawlak & Vicente, 1996). The P interface displays primarily physical information about the work domain. In contrast, the P+F interface

(designed under EID principles) displays in a cognitive relevant manner both physical and functional information about the work domain by means of configural displays.

As the magnitude of sensor noise is increased on both displays, we expect performance to worsen and stability to decrease for both P+F and P operators, even though we anticipate that performance of P+F operators will not be inferior than that of P operators. Control strategies are also expected to change while operators learn how to cope with the noise, outlining aspects of adaptation to the uncertainty in the environment.

Our research is expected to contribute to the EID literature. It will be the first study to examine the impact of sensor noise on the robustness of EID by manipulating the magnitude of sensor noise. Moreover, it will also be the first study to address the issue of adaptation to uncertainty in the environment by manipulating sensor noise in a process control microworld.

Finger, R., & Bisantz, A. M. (2002). Utilizing graphical formats to convey uncertainty in a decision-making task. *Theoretical Issues in Ergonomics Science*, 1, 1-25.

Johnson, C. D. (1997). *Process control instrumentation technology, Fifth Edition*. Upper Saddle River, NJ: Prentice-Hall Inc.

Pawlak, W., and Vicente, K. J. (1996). Inducing effective operator control through ecological interface design. *International Journal of Human-Computer Studies*, 44, 653-688.

Stein, P. K. (1969) The response of transducers to their environment: The problem of signal and noise. *Lf/MSE Publication*, 17, 1-15.

Vicente, K. J., and Rasmussen, J. (1990). The ecology of human-machine systems II: Mediating "direct perception" in complex work domains. *Ecological Psychology*, 2, 207-250.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems - Man and Cybernetics*, 22, 589-606.

Woods, D. D. (1988). Coping with complexity: The psychology of human behavior in complex system. In L. P. Goodstein, H. B. Anderson, and S. E. Olsen (Eds.), *Tasks, errors, and mental models: A festschrift to celebrate the 60th birthday of professor Jens Rasmussen* (pp. 128-148). London, UK: Taylor and Francis.

Ecological Analysis and Judgment Analysis

Tom Stewart

University at Albany, SUNY, USA
t.stewart@albany.edu

At the last year's meeting, I tried to differentiate "judgment analysis" and "policy capturing". There is a note on this distinction on the Brunswick Society website (<http://www.brunswik.org/notes/japc.html>). I just read it, and I recommend it. It includes comments from Joe Ward, Jim Naylor, and Ray Cooksey.

The main point is that judgment analysis requires more than just the use of regression analysis for modeling judgments. It requires both idiographic data analysis and ecological analysis, that is, detailed analysis of the task and careful attention to representative design. Policy capturing refers to single system studies, usually with artificial tasks, where little attention is given to ecological analysis, and sometimes judgments are averaged over subjects before the regression is done (nomothetic analysis). Policy capturing may be useful in certain situations, such as group facilitation, where generality is not required, but it is inappropriate to cite Brunswick or Hammond or Cooksey as support for a policy capturing study. Both the term and the method were invented by others.

What does adequate ecological analysis for a judgment study look like? Ideally, it involves a combination of methods for assessing the properties of the judge's environment. Interview, observation, think-aloud protocols, and analysis of existing data all have a role. This takes time and effort, but is essential.

Task sensitivity analysis is a technical component of ecological analysis. It addresses two closely related questions: 1) Is accuracy sensitive to importance weights? and 2) How precisely can weights be estimated? To answer the first question, we simulate a large number of judges who use random, but correctly signed, weights, and we examine the resulting distribution of G. An Excel add-in called Crystal Ball handles this simulation easily. Often we will find that G is high regardless of the weights used. Then we have to ask "If the judge is going to be accurate no matter how he or she weights the cues, how can learning occur?"

To answer the second question, assuming a reasonable range of values for R^2 , the standard formula for the standard error of a regression coefficient applies. If you are uncomfortable with the assumptions of the standard error formula, simulation or statistical bootstrapping could be used. If non-linearity or non-additivity is to be investigated, the power to detect them can also be examined, but we

already know that the typical judgment analysis design will not have sufficient power to detect these effects. I can't imagine proceeding with a judgment analysis study without knowing the potential accuracy of parameter estimates. Consideration should be given to improving those estimates by applying Gary McClelland's suggestions for improving the efficiency of representative designs (<http://www.brunswik.org/notes/essay5/essay5.html>). I plan to write a paper on methods for task sensitivity analysis.

I have just done a sensitivity analysis on a task based on real patients. There were 50 cases and 6 cues, with very low cue intercorrelations. The variability in the accuracy correlation over 10,000 simulated judges with random weights was less than the estimated sampling variability for that correlation. Obviously, this task will not reliably identify differences in accuracy stemming from weights.

I hope that in conducting our research and in reviewing papers and research proposals involving judgment analysis, we will pay close attention to the ecological analysis. It may turn out to be more important than the regression results.

Logistic Lens Model analysis in Medicine

Thomas G. Tape

University of Nebraska College of Medicine, USA
tgtape@unmc.edu

My colleagues and I have been working on applications of the logistic lens model in medicine. Most medical tasks involve dichotomous judgments such as whether or not a disease is present. Such tasks do not easily lend themselves to the traditional linear lens model approach. We usually worked around the problem by asking physicians to estimate the probability of disease and then use the probability estimates to build a linear judgment model. Models of the environment for many medical conditions have been published but nearly all are based on logistic regression. We had to violate assumptions of the models to make our data fit.

Tom Stewart worked out the mathematics of the logistic lens model equation some time ago and the details described by Ray Cooksey in his Judgment Analysis book. Although the methods have been published, we have not been able to find any applications of logistic lens model analysis in the medical literature. Thus we decided to try it on two of our own datasets.

The first study looked at medical student diagnoses of meningitis based on paper vignettes of real cases. We obtained both dichotomous diagnoses as well as

probability estimates. This allowed us to compare the traditional lens model approach with the probability estimates to the new logistic lens model of the categorical judgments. We found the two approaches to be virtually equivalent in their ability to model achievement.

Our second study involved practicing physicians caring for women with suspected urinary tract infections. The practitioners prospectively recorded the patients' signs and symptoms along with their diagnostic assessment before any laboratory testing. We then did urinalysis and urine cultures to determine the actual diagnosis. When looking at models based on signs and symptoms alone, the value of C was rather large. This led us to wonder if we had left out an important cue the subjects were using.

When we added the results of office lab testing (which they were not supposed to look at until after making their initial diagnosis), the value of C dropped dramatically. We suspect that our efforts to obtain diagnoses unbiased by office laboratory results failed.

We will present the details of both studies at the Vancouver Brunswick meeting. We are encouraged by our preliminary exploration of the logistic lens model equation and look forward to learning more about it.

A Graphical Artifact by Brunswick

Ryan D. Tweney

Bowling Green State University, USA
tweney@bgsu.edu

Our work on Michael Faraday's "epistemic artifacts" continues, and we are now trying to organize the 1000+ digital images of his gold and metallic specimens, and to integrate these with his diary for 1856; A long slow project, but continually exciting as we see more and more of the "forest" behind all those trees!

As something of a breather for me, Mike Doherty and I have had several conversations about Egon Brunswick's wonderful Figure 32 (p. 90 of *Perception and the representative design of psychological experiments*, 1956). For me, it is a terrific example of how much information can be conveyed by the right graphical representation. Anyone who has felt that Brunswick's own writings are "difficult" should study this figure. Whatever the merits of his prose style, this figure conveys so much that it is clear to us that Brunswick was indeed a masterful expositor!

We'll do more with this, particularly in the context of our talk at the forthcoming Leuven meeting next summer. Meanwhile, I'm sorry not to have included it

in a forthcoming book chapter on a related topic involving "cognitive artifacts," an historical paper on the decline of psychology's use of "brass and glass" instruments, and their replacement by statistical "instruments." This is to appear in an edited volume by David Baker, due out from the University of Akron Press any day now.

Modeling and Mentoring

Elise Weaver

Worcester Polytechnic Institute, USA
eweaver@WPI.EDU

My work over the past year has been heavily influenced by my teaching activity at Worcester Polytechnic Institute (www.wpi.edu). WPI specializes in training technology-oriented students, and the traditional role of our Social Science and Policy Studies department has been to mentor these students through a junior-year society and technology project. More recently, our department has taken on a focus on system dynamics, a computer simulation technique rooted in control theory, and I was hired to be a psychologist and system dynamicist. To this end, I coordinated a symposium within the system dynamics conference on psychology, in which Alex Kirlik, John Flach and Jim Townsend presented and John Sterman was discussant.

Because system dynamics has to do with simulating behavior over time, I have been thinking (even more than usual) about how judgment importance weights change over time. I have long been thinking about representing policy for judgment as a position vector of importance weights in a multidimensional policy space. It seems to me that one's position about what is important can be represented mathematically and can be changed through persuasion or learning according to certain rules, like the Rescorla-Wagner equation of learning theory. This is a theoretical framework, but I think it could yield a rich set of empirical questions such as what kinds of influences force one out of one's prior position? Are the rules for motion in response to evidence Bayesian, not in the "log odds" sense, but in the sense that prior uncertainty in one's importance weights (i.e. an open mind) yields more flexibility for motion to a new policy position? I will present this framework at our Brunswick Society meeting in November.

Another project I have engaged in with George Richardson of University at Albany, SUNY and Kent Rissmiller of WPI, is a system dynamics simulation of a speculation of Ken Hammond's about how policy thresholds in a Taylor Russell diagram would move in response to stakeholder pressures. Because of this work, I have been invited to participate in an NRC-

sponsored workshop on the development of performance based building code standards.

This work has been extended this year to a model combining Ken's idea with ideas from Cole & Dempsey's (2002) book about the use of guilt by association as a decision rule in security investigations. These authors make the case that a poor decision rule not only compromises civil liberties, but also compromises the quality of the investigation. Our system dynamics model with the oscillating threshold converges to a better judgment rule over time to minimize false positive and false negative cases when there are feedback mechanisms for stakeholder influence.

My work continues with Tom Stewart on an empirically derived map of judgment competencies. I am currently learning the conceptual distinction between a cognitive ability and a judgment competency, a distinction necessary to eliciting the appropriate set of reviewers for our work!

Finally, I will be living in D.C. for 7 weeks, as I am involved in mentoring five junior-year society and technology team projects in government agencies. These groups will be consulting to NOAA about aquaculture and the environment, NSF about undergraduate education in nanotechnology, the US Patent Office in arranging a curriculum to train their patent examiners in recent computer technology, the Consumer Product and Safety Commission to advise on educational and technological interventions in residential pool safety, and to Montgomery County to advise on the travel patterns of their ridership.

Communication, Judgment and Decision Making in Medicine

Robert S. Wigton

University of Nebraska Medical Center College of
Medicine, USA
Wigton@Unmc.edu

I am working on two projects to study the outpatient diagnosis and management of respiratory tract infections (cold, flu, bronchitis, sore throat, sinusitis). The diagnosis of common respiratory infections is important because the pervasive overuse of antibiotics to treat mostly viral illnesses has led to a dangerous increase in the number of antibiotic resistant bacteria, reducing our ability to cure more dangerous life threatening infections.

The first project, working with Tom Tape and Gay Canaris, is a study of 327 patients who came to our clinics complaining of respiratory symptoms over a 3 year study period. Both practitioners and patients filled out an extensive symptom questionnaire and practitioners also reported physical findings. Both

groups recorded what they thought was the cause of the illness. We found only fair agreement between practitioners and patients regarding what symptoms were present. We found considerable disagreement about the cause of the illness and differences in symptom weighting. These findings suggest approaches to improve communication and possibly diagnostic accuracy.

In a second study, working with Ralph Gonzales and Carol Darr, we conducted case vignette studies with 101 Denver practitioners on two occasions: before and one year after an intervention designed to decrease antibiotic use. The goals were first to determine the baseline policies and then to measure whether the practitioners were influenced by patient factors, such as demand for antibiotics, in addition to symptoms and exam findings. I also am examining whether changes in prescribing behavior in actual patients are accompanied by changes in cue weights and in prescribing behavior in the paper cases. In addition, I sent out a subsequent study to all respondents to see if they can recognize their own policies. At the Brunswick meeting, I'll present the preliminary results of the policy capturing and of the comparison of each individual's policies in year 1 and year 2. These results show the measured policies are quite stable from year to year.

Brunswikian Representative Design in the Research of Persistence in Educational Psychology

Bernhard Wolf

University of Landau, Germany
wolf@uni-landau.de

In a longitudinal design, German children are evaluated three times concerning the motivational characteristic "persistence" with respect to the achievement domain. The first time these children were six year old and in their last year of nursery school (1997). By the second time of the evaluation, the same children were pupils at the end of primary school and about 10 years of age (2001). The design will finally be completed in 2006, at which point these pupils will be 15 years old, attending secondary school.

In a Representative Design (Brunswick) complex individual scores of some single cases will be computed using information of a great amount of situational items of persistence for each case. The samples of situations are focused in such a design, not the sample of individuals. These situation-based indicators are combined over all three times of the design generating a measure which will be compared with achievement at school of that special pupil. At all

of these three stages the indicators of persistence are combined in a kind of situational array. Within each of the three stages graded weighting procedures will be used for the analysis. Natural behavior and everyday life are pointing the way for that kind of research.

The combination of single case approach with forms of representative design facilitates the comprehension of psychological processes between persistence and school-achievement.

Clear Logic and Fuzzy Guidance: A Policy Capturing Study of Merit Raise Decisions

Kenneth M. York and Lizabeth A. Barclay

School of Business Administration
Oakland University, USA
york@oakland.edu

Merit pay is widely accepted as a way to recognize and reward performance, but merit pay systems are problematic. For merit pay systems to be effective, organizations must consider the relationship of such systems to business objectives, so that the organization is rewarding what it intends to reward. This study used policy capturing to determine the weights an actual merit raise committee gave to different sources of information, and compared the results to the policy the merit committee was charged with implementing. Actual performance data, ratings, and merit pay increases for a set of employees is often unobtainable; a Freedom of Information Act (FOIA) request was used to obtain a Merit Committee's worksheets, annual reports, and the actual percentage increases given to a sample of 36 business school faculty. The charge the faculty gave to the Merit Committee was that the committee "shall consider each faculty member's contributions in research, teaching, and service, with special consideration for research." The Merit Committee was very consistent in their judgments ($R = .90$), and gave the most weight to research. As a percentage of the total unique variance accounted for, research accounted for 80%, teaching for 19%, and service for 1%.

A second policy capturing analysis was done to determine whether the Merit Committee's judgments could be captured directly from the data they used to make their judgments. Instead of using the Merit Committee's ratings for research, teaching, and service, tallies were made of items listed in the annual reports under the categories of research, teaching, and service. For this regression analysis, research was operationally defined as the total number of publications and presentations; teaching

as the mean rating across six items on the student course evaluation form; and service was the number of department, school, university, and professional service activities listed on the annual report. Using this "objective" data, the policy of the Merit Committee was captured ($R = .67$), and the percentage of unique variance accounted for showed a heavy emphasis on research. Based on these results, the short-term strategy for maximizing merit pay increases would be to put effort into research and reduce the amount of investment in service and teaching, which is counter to the organization's stated objectives of rewarding research, teaching, and service. In this case, was the intent to reward research, teaching, and service undermined by decisions made by a merit pay committee which overwhelmingly rewarded research—another example of the folly of rewarding A while hoping for B?

York, K. M., and Barclay, L. A. (2003). Clear logic and fuzzy guidance: A policy capturing study of merit raise decisions. *Public Personnel Management, 32*(2), 287-299.

York, K. M., and Barclay, L. A. (2003). Clear logic and fuzzy guidance: A policy capturing study of merit raise decisions. *Public Personnel Management, 32*(2), 287-299.

Tentative Agenda

Nineteenth Annual International Meeting of the Brunswik Society 100th Anniversary of Brunswik's Birth Vancouver, British Columbia, Canada

November 6-7, 2003

Cypress Room, Hyatt Regency Vancouver

Thursday, 6 November

- 13.00 - 13.30 – **Late registration**
- 13.30 – 13.45 – **Introductions and Welcome**
(Jim Holzworth, Mandeep Dhama, Elise Weaver, Tom Stewart)
- 13.45 – 15.00 – **Paper session 1: Medical Judgment and Decision Making**
(Chair: Neal Dawson)
Robert S. Wigton, C. A. Darr, K. K. Corbett, D. Nickol, & R. Gonzales –
How stable are practitioner's strategies from year to year in
deciding to prescribe antibiotics for respiratory tract infection?
Thomas G. Tape - Applications of the Logistic Lens Model
Liz Smith, Ken Gilhooly, & Anne Walker –
Clustered or flustered: Prescribing of antidepressants in primary
care
- 15.00 – 15.15 – **Tea and coffee break**
- 15.15 – 16.15 – **Discussion session 1: What impact can and should
Brunswikian research have on challenging and changing
policy and practice?** (Chair – Marcus O'Connor)
Discussants – Jeryl Mumpower and Tom Stewart
- 16.15 – 17.30 – **Paper session 2: Judgment Models** (Chair: Bernhard Wolf)
John O. Ekore –
Managers' self-efficacy differences and choice of decision-
making style in banks
Robert M. Hamm –
Linear versus logistic lens models for predicting men's prostate
cancer screening
Grant Broad, Mandeep K. Dhama, & Peter Ayton – Simple heuristics and
regression models of predictions of violent offending
- 17.30 – **Adjourn**
- 18.30 – **Evening Group Dinner at a restaurant in Vancouver**

Friday, 7 November

08.30 – 09.00 – **Tea and coffee break**

09.00 – 10.35 – **Paper session 3: Theory, Methods, and Analysis**

(Chair: Tom Stewart)

Bernhard Wolf –

Fritz Heider and Egon Brunswik: Their lens models: Origins, similarities, discrepancies

Kathy Mosier –

Coherence and correspondence: Complementarity in the context of technological advances.

Konstantinos V. Katsikopoulos –

How fast and frugal can the lens model be?

Elise Weaver -

Geometrically G: An investigation of movement through a judgment policy space

10.35 – 10.50 – **Tea and coffee break**

10.50 – 12.00 – **Discussion session 2: Coherence and correspondence approaches to J/DM: How can we do more to integrate them?** (Chair – Jim Holzworth)

Discussants – Kathy Mosier, Mike Doherty, and Ken Hammond

12.00 – 14.00 **Buffet lunch followed by an invited presentation by Ken Hammond on ‘The Next 100 years (minus 80).’**

14.00 – 15.40 – **Paper Session 4: Research on Experts-Novices and New Topics** (Chair – Clare Harries)

Priscilla Harries –

Improving clinical judgments using captured policy

James Shanteau –

Development of expertise in air traffic control

Phil T. Dunwoody, A. Goodie, & R. P. Mahan –

A Brunswikian look at classic base rate neglect

David J. Weiss, & Christian Schunn –

Evaluating peer evaluation of writing

15.40 – 16.00 – **Tea and coffee break**

16.00 – 17.00 – **Discussion session 3: What has been the impact of new technologies on Brunswikian research?**

(Chair – Chris Anderson)

Discussants – Alex Wearing and Gary McClelland

17.00 – 17.15 – **Brunswik-Hammond New Investigator Prize (Awarded by Ken Hammond)**

17.15 – **Farewell and Meeting adjourned**