



Edited by Clare Harries

October 2001

Good Read: *The Essential Brunswik*

I received my copy of "The Essential Brunswik" this week. Let us all express our thanks, appreciation and congratulations to Ken Hammond and Tom Stewart for such a magnificent job. And let's express our thanks and appreciation to Oxford University Press.

I think that the book is a masterwork.

Mike Doherty

Seventeenth Annual Meeting of the Brunswik Society - See page 20-21

Research news

2	Adelman	Testing the Effectiveness of Icons for Supporting Distributed Team Decision Making Under Time Pressure
2	Albright	Interpersonal Perception
2	Alvarez	A contextual-normative approximation to the integration of multiple cues in learning
3	Athanasou	The Role of Cue intercorrelations in judgments
4	Cooksey	Expanding the Scope of the Brunswikian Perspective
4	Cvetkovich	Processing information about trust in the deep social mind.
5	DeKay	A Brunswikian investigation of risk ranking
6	Denig	Analysis of GPs' prescribing decisions
6	Dhami	Ecological Rationality
6	Earle	Testing the dual-mode model of cooperation
8	Funder	Personality judgment and the Realistic Accuracy Model
8	Gilhooly, Smith, & Harries (Cilla)	Referral prioritisation in Occupational Therapy and Guideline use in General Practice
8	Hamm	Patient Education
9	Hammond	Environmental structures, Cognitive structures and the Information Age.
10	Harries (Clare)	Strategic and Medical Decision Making; What, when and how
10	Holzworth	Brunswikian research at the University of Connecticut
11	Jha & Bisantz	Modeling fault diagnosis in a dynamic process control task using a multivariate lens model
12	Joyce	Judgment Analysis of Quality of Life Assessments
12	Juslin (Patrik)	Cognitive Feedback and Musical Expressivity: Teaching the Thing that Cannot be Taught
13	Juslin (Peter)	Process and representation in multiple-cue probability learning
13	Koele	The continuing search for configularity
13	McClelland & Fasolo	Representative Design Revisited
14	Mosier	Correspondence, coherence and the cognitive continuum
15	Mumpower	Intellectual activities of a Brunswikian administrator
15	Nath	Analyzing the utility of utilities and conceptual disease spaces
16	Newell	Fast and frugal: Some of the people, some of the time, some of the environments?
16	Shanteau	Evaluating Expertise
18	Shim	Exploration of Alternative Designs for Judgment Analysis Application in Public Policy Formulation
19	Tape	Modeling and predicting cardiac disease in "low risk" patients with chest pain
19	Tweney	Brunswikian accounts of scientific thinking
20	Weaver & Stewart	Why are some judges better than others?

Testing the Effectiveness of Icons for Supporting Distributed Team Decision Making Under Time Pressure

Leonard Adelman, Sheryl L. Miller, and Cedric Yeo
(George Mason University, USA)

The research we described at Brunswik 2000 showed that the flow of information among distributed team members decreased with increasing time pressure, with a corresponding decrease in decision accuracy (achievement, r_a). The study performed this year tested the effectiveness of a "Send" icon to remind team members to send information upon its receipt, and a "Receive" icon to tell them when they had new information, in a simulated team decision making task. As predicted, the "Send" icon was effective in maintaining information flow, particularly under higher levels of time pressure and when simulated teammates sent less information. However, contrary to our prediction, the "Receive" icon was not effective under the highest time pressure level. This occurred because participants' using the "Receive" icon adopted a strategy of making decisions before, not after receiving information as time pressure increased; as a result, they made a greater proportion of decisions with less information, and with less cognitive control (R_s) under the highest time pressure level. The study also showed that time pressure's general, negative effect on accuracy was not caused by participants adopting new strategies for combining information into decisions (G) but, rather, because of lower cognitive control (R_s) in applying the strategy they were trained to use, regardless of how much information they had. Conceptually, the results illustrate the close and sometimes subtle relationship between the task, displays, strategy, and achievement. In addition, the study found that participants with higher working memory capacity were better able to integrate more information and that task variables (time pressure, amount of information, run number, scenario order, and type of information) had strong effects on all performance variables.

Leonard Adelman: ladelman@gmuvox.gmu.edu

Interpersonal Perception

Linda Albright (Westfield State College, USA)

My research continues to center around interpersonal perception. In a recent study we consider interpersonal perception along with personality, interpersonal behavior and interpersonal relations. In groups of four individual participants in this study interacted separately with two individuals from his or her group and then the group interacted as a whole. Some groups consisted of people who

were previously unacquainted, whereas others consisted of people who were highly acquainted. We videotaped both the dyadic and the group interactions and coded various nonverbal and verbal behavior. In addition participants judged themselves and each other on a number of dimensions following the interactions, as well as relational aspects of the interaction itself. Measures of personality were obtained prior to the interactions. Some Brunswikian questions we will address how personality relates to interpersonal behavior and how behavior relates to interpersonal perception. Other questions pertain to the "chemistry" of social relations: can the relational aspects of the interaction be predicted by various personality dimensions. What interpersonal behaviors underlie positive interactions? We will also test some hypotheses about individual differences in judgmental accuracy.

In another recent study, we examined interpersonal perception in the context of computer-mediated communication (CMC). These data have been collected, but we have not begun any analysis. Participants in this study were randomly assigned to groups of four unacquainted individuals and were requested to exchange emails with two others persons on a weekly basis for six weeks. After each email exchange participants rated themselves and each other on various dimensions and rated the various aspects of their relationship. Participants also reported metaperceptions (inferences of how they were judged by their partners). Content analyses of the email messages will be performed. How well individuals get to know each other through CMC and how interpersonal judgments are related to the nature of their verbal exchanges are some questions we will pursue.

Linda Albright: l_albright@foma.wsc.ma.edu

A Contextual-Normative Approximation to the Integration of Multiple Cues in Learning.

Manuel Miguel Ramos Álvarez (University of Jaen, Spain)

The main result obtained in our laboratory up to now is that, in blocking procedures, the processing of predicting cues does not follow the competitive mechanism, that is to say, it cannot be explained with the discounting or competition principle.

Throughout this year we have perfected the Integration formal model previously proposed (see Article #17, 2000). Following the lens model, the ecological versus subjective regression, we have focused on the subjective system. We have incorporated a regression mechanism based on a normative contextualization principle. This formal mechanism determines whether the predictive validity of a specific event will be

computed in a relative (for example, with a competition principle in mind) or independent way. We basically estimate, through conditional probabilities, if the previously mentioned event will introduce predictive changes in connection with other alternative events. Besides, such a mechanism is dependent upon a series of beliefs relevant to the interaction. We have carried out an experimental series where we study the effect of Conceptual processing on the predictive competition. By means of careful written instructions, we introduced specific biases dealing with each predictor (experiment 2), about the predictors' relationship (experiment 2) or in connection with an independent versus dependent strategy. The methodology we used was better than other commonly used methodologies for our contextual proposal, as the simulations we carried out have proven. Results show that the relativization principle does not systematically intervene in any situation including complex stimuli. This series of experiments makes some beliefs very relevant, such as the idea that there is some previous probability to the shadowed stimulus or the view that predictors have a strong connection between themselves. These facts have led us to conclude against current models and in favour of contextualization model.

In addition, we have worked and perfected several computer programs related to experimentation and modelization. First, the "Judgment" program allows controlling and carrying out experiments dealing with every type of causal-contingency research, including the Cue Probability Learning Paradigm. Second, the "Integra2000" program allows the testing of our formal model by making a computation of the most relevant indexes in the Lens Model (i.e. consistency, etc.). Third, the "AsocJudg" program is able to simulate the most relevant associative models, the basic Lineal Operator model (Rescorla and Wagner, 1972), The net model that codes information absences of stimuli (Tassoni, 1995), The Configural model of conditioning (Pearce, 1987, 1994) and soon we will also be able to simulate a new connectionist model, the RASHNL one (Kruschke and Johansen, 1999). Fourth, the "NormJudg" allows simulating the inductive models, the Probabilistic Focal model (Cheng and Novick, 1992) and Power-PC one (Cheng, 1997).

Manuel Miguel Ramos Álvarez: mramos@ujaen.es

Role of Cue Intercorrelations in Judgments

James A. Athanasou (University of Technology, Sydney, Australia)

This research program has been considering the role of cue intercorrelations in the single lens model. A graduate student, Olu Aiyewalehinmi, and I have

varied the cue intercorrelations from 0.0 through 0.3, 0.6 to 0.9 and looked at the effects on the lens model parameters.

As part of this research program we used 75 profiles containing six cues rated from 0 to 9. The cues described different aspects of a course or subject. The six cues used were: how easy the subject was; the quality of teaching; the importance of the subject; whether it was one's best subject; the amount of time spent studying; and the amount of time spent on assignments, projects etc. These cues were selected on the basis of earlier lens model analyses. People were asked to decide how interested they would be in studying the course or subject that was described.

Fifteen out of the 75 profiles were repeated to indicate the test-retest reliability of decisions. The second of four studies was conducted using university students in Nigeria and we noted very low test-retest reliabilities for the judgements. These did not vary in conjunction with the extent of cue intercorrelation.

Also, the multiple correlations between cues and judgements did not increase with the extent of cue intercorrelation as much as in the study reported last year in this newsletter. Multiple correlations ranged from 0.21 to 0.60 for the cues that were intercorrelated 0.0; 0.22 to 0.66 for cue intercorrelations of 0.3; 0.23 to 0.72 for cue intercorrelations of 0.6; and 0.13 to 0.99 for cue intercorrelations of 0.9.

One question that has arisen from this program of research relates to ensuring the quality of the data. Despite monetary rewards for participation, it was not clear that reliable and valid results were obtained. A second issue to resolve is the quantity of judgements required in a single lens model study in order for decision parameters to stabilise. Other features of the decision that need to be considered in future investigations include the number of cues, the validity of the rating scale and cue content. We are now moving towards examining the effect of other cue contents on judgements within the single lens model.

Copies of some earlier reports of lens model studies in the judgement of interest include:

Athanasou, J A. (1999). Judgements of interest in vocational education subjects. *Australian and New Zealand Journal of Vocational Education Research*, 60-76

Athanasou, J A. (1998). Perceptions of interest: A lens model analysis, *Australian Psychologist*, 33, 223-227.

Athanasou, J A & Aiyewalehinmi, O. (2001). A case study of student judgements of interest in a subject, *Ifa Psychologia*, 9(1), 74-81.

Athanasou, J A & Cooksey, R W (in press)
Judgment of factors influencing interest: An
Australian study. *Journal of Vocational Education
Research*.

Finally, a brief editorial on Brunswikian approaches
that may have relevance for career development
research will appear in the *Australian Journal of
Career Development*, volume 10 (3), 2001.

James Athanasou: Jim.Athanasou@uts.edu.au

Expanding the Scope of the Brunswikian Perspective

Ray W. Cooksey, (University of New England,
Armidale, Australia)

I have two projects to report on this year. The first
relates to a Commonwealth Government-funded
collaborative research project with colleagues in
Queensland, Australia, where we are employing a
Brunswikian approach to the study of teacher
judgments of Year 5 students' levels of written
literacy. The study has combined judgment analysis,
cause mapping methods and think-aloud protocols
to investigate how teachers make their judgments of
literacy, when confronted with specific exemplars of
students' written work in a particular genre. Teachers
work through a series of pieces written by
their own students (judgments in context) and a
series of pieces sampled from a wider group of
students studying with other teachers (judgments
out of context). Teachers think aloud as they make
their judgments and all talk is recorded. Teachers
are also asked to construct cause maps of the
factors they see as being causally connected to
students' literacy achievements. Currently, the
cause maps are being analysed and compared for
conceptual density, interconnectedness, and
thematic variation. Think-aloud protocols are being
used, in conjunction with teachers' judgments of
achievement, to produce an enhanced cue set that
will model not only the influence of physical, textual
and genre writing features, but also the intrusions
and importations of other cues on a case by case
basis (knowledge of a student's home background,
gender, prior achievement and ability, etc). In other
words, we are trying to capture the dynamics of the
cue ecology as it gets mapped onto teacher
judgments in situ and trying to understand the
implications of making judgments out of context (as
is done on standardised tests of literacy) versus
making judgments in context (as is done in the
classroom where teachers know the producers of
the works they judge) have for judgment accuracy
and complexity. The cue set is currently being coded
for the judgment analysis stage.

The second project has resulted in a very recent
publication in the new Lawrence Erlbaum organisation
science journal *Emergence* (2001, vol. 3, no. 1)
entitled 'What is Complexity Science? A Contextually-
Grounded Tapestry of Systemic Dynamism, Paradigm
Diversity, Theoretical Eclecticism, and Organizational
Learning'. In this article, I explore the implications of
the perspectives of Brunswik, Churchman, and other
systems and complexity theorists to argue for a new
generalist perspective on behaviour research having
an organisational focus. For me, this is what I have
always thought Brunswik's work could ultimately
contribute to – a rethinking of the science of the study
of human behaviour in a way that values paradigmatic,
theoretical and methodological diversity, that is, itself,
a learning science, that values systemic, dynamic and
nonlinear thinking, that is contextually grounded and
that avails itself of the full armoury of tools and
techniques that researchers can bring to bear on
understanding a particular problem, be it theoretical or
practical. While this is not decision research per se, I
see strong implications for decision research in the
new perspective; members of the Brunswik Society
may remember seeing snippets of this evolving
perspective peeking through some of the papers I
presented at Society meetings in the mid-1990's as
well as in the last chapter of my book on judgment
analysis. There are also some important implications
for the diversification of education and training for
researchers and practitioners in the behavioural
sciences. For me, this complexification of the human
activity of behavioural research has emerged as my
new playground – sorry, Ken, you once tried to warn
me not to follow this path, but I did anyway! For any
who are interested, I can forward an electronic copy of
the paper to you.

Ray Cooksey: rcooksey@metz.une.edu.au

Processing Social Information

George Cvetkovich (Western Washington University,
USA)

I've continued work based on the Salient Values
Similarity (SVS) model of trust begun with Tim Earle.
According to the model the attribution that one shares
important (salient) values with another involves
information processing along the cognitive continuum.
Brain-imaging studies have now identified the location
of some of the more important unconscious, automatic
processing tasks related to trust (e.g., basic abilities to
mentalize or "mind read"). This mode of processing
combines with conscious, rational processing to
integrate implicit and explicit understandings of how
the human mind works with available information about
the actions, statements, emotional states, etc. of a
particular individual. The SVS model was explored in
several recent studies with Michael Siegrist (Univ. of
Zurich) and others that examined risk and benefit

perceptions of hazardous technology, the relative importance of trust when one is knowledgeable about a topic or not, the willingness to accept chance as an explanation for cancer clusters, and the effects of new information on changing trust. Pat Winter (US Forest Service) and I have just completed a study indicating that conflicts over the management for the protection of endangered species in the national forests result from four patterns of social construction and representation of value saliency and similarity. I will soon be returning to working on a book, "The Trick of Trust: Evolution, Personal Preferences, Social Issues and the Deep Social Mind," aimed at a scientifically literate, but non-professional audience, utilizing the SVS model to examine trust in various domains (interpersonal relationships, politics, health care, etc).

George Cvetkovich: George.Cvetkovich@wwu.edu

A Brunswikian Investigation of Risk Ranking

Mike DeKay (Carnegie Mellon University, USA)

Since coming to Carnegie Mellon in 1996, I've been involved in a large NSF/EPA project on ranking risks that has something of a Brunswikian flavor at times. The project has involved several other faculty (Granger Morgan, Baruch Fischhoff, Paul Fischbeck, and Keith Florig) and graduate students (Karen Jenni, Kara Morgan, Jun Long, Claire Palmgren, and Henry Willis). I'll talk about some of this research (which part I don't yet know) at the conference (see back page of this newsletter – Ed.), but here is a summary of the high points.

In the most general terms, the project is designed to improve on the methods used in the many national, state, and local risk-ranking efforts that have been undertaken or funded by EPA over the last 10-15 years.

One of the first things you have to decide when you rank risks is what to call a risk. In other words, how should risks be sliced or lumped into categories, and what are the consequences of these decisions? Our thoughts (which I won't go into here) are summarized in a paper that appeared in *Risk Analysis* last year.

We also needed a case or situation in which to test our ideas (without the strings and pressures of the real world), so we developed a test bed involving the risks to students in a hypothetical middle school. We developed risk-summary sheets for 22 risks in this context, with each risk described in terms of a common set of 12 attributes, and with text that discusses the risk in general and at the school. A paper that summarizes the development of this test bed, the choices of attributes, and the procedures

for having individuals and groups rank the risks using these materials is now in press at *Risk Analysis*. There is also a book chapter that appeared this summer in a "Resources for the Future" volume.

We have compared the rankings that result when individuals use either the full summary sheets, a brief text description only, the attribute table only (with generic titles like Risk A, Risk B), or the combination of the text and table (but omitting additional information from the full summary sheets). We compare agreement (measured as the mean pair-wise correlation among rankings) both within and between materials conditions, and also compare rankings to the ranking by expected mortality. We have done this study both with college students and with lay people, but we have not yet written up the paper.

One of the key metrics in our ranking studies is the relationship between holistic rankings and multi-attribute rankings. We have experimented with several models for specifying the single-attribute utility functions and the attribute weighting functions for constructing the multi-attribute rankings. Results suggest that in some cases, simpler models that do not require time-consuming elicitation of functions from participants may perform just as well as more traditional models, at least when performance is measured as the convergence between holistic and multi-attribute rankings. A paper is in the works.

We have also run many group risk-ranking exercises, some involving risk professionals (at a Harvard short course) and some involving lay people. A paper based on over 40 groups from Harvard is in press in *Risk Analysis*. Among other things, we assess whether the holistic and multi-attribute models tell similar stories (they do, both at the individual and group levels), how agreement among individuals changes as a function of the group discussion (it increases), and whether individuals buy into their groups' processes and rankings (they do, as judged by self-reports and regression-based measures). A paper is currently in press at *Risk Analysis*.

With layperson groups, we have investigated the effects of the presence and order of the tasks in the ranking exercise, with particular attention to the holistic and multi-attribute ranking procedures. Analyses have not been completed, but the current story is that our standard ordering (holistic, then multi-attribute, then compare and revise) works quite well, and there are tangible benefits to including the multi-attribute approach along with the holistic approach. A paper should result.

Most recently, we have been expanding the test bed to include ecological risks and their attributes (we are now using the entire county rather than the school). It is a major challenge to select the ecological attributes

in an informed way, and we have conducted a number of studies on this front. Some of these are traditional factor-analytic studies with group-level data, whereas some allow for factor analysis and multi-dimensional scaling at the individual-subject level. A recent effort takes this individual-level approach to compare perceptions of laypersons, environmentalists, industry ecologists, and government ecologists. Finally, we have just collected data from four Harvard groups that ranked ten of the new risks, using the new materials and attributes. Papers should result, but they are not first in the queue.

In my view, this research is somewhat Brunswikian in that we have devoted a great deal of attention to the environment and stimulus set, we have allowed attribute values to vary and correlate across risks (although we have not randomly chosen the risks, in part because it's not entirely clear how one would do that), we have incorporated metrics and analyses (e.g., policy capturing, not mentioned above) that are familiar to Brunswikians, and we have been careful to analyze the data at the individual-subject level when possible.

Mike DeKay: deKay@andrew.cmu.edu

Analysis of GPs' Prescribing Decisions

Petra Denig (Groningen, The Netherlands).

At the Department of Clinical Pharmacology, the Rational Drug Use group focuses on improving prescribing behaviour of doctors. We have used clinical judgement analysis as method to analyse prescribing decisions and as an educational tool in group meetings with general practitioners (GPs). As mentioned previously, we participated in the Drug Education Project in which the effect of an educational program including cognitive feedback (based on CJA) was tested in four countries (Veninga CCM et al. in the Am J Resp Crit Care Med 1999;160:1254-62). The experiments we conducted in the Netherlands were described in more detail by Veninga CCM, Denig P, Zwaagstra R, Haaijer-Ruskamp FM. Improving drug treatment in general practice. J Clinical Epidemiology 2000; 53:762-72. Our conclusion in the Netherlands was that discussing a combination of outcome feedback, cognitive feedback and optimal decision policies was effective in improving actual prescribing behaviour for asthma and urinary tract infections. Following these experiments we wrote a review of studies using cognitive feedback to improve the quality of treatment decision of doctors. This will hopefully be published in *Medical Education* in the beginning of next year. Some issues discussed in this paper are: (1) The value of cognitive feedback in continuing medical education is limited to tasks and problems

that can be adequately described on paper. (2) In a learning setting, cognitive feedback provides reflection which is based on systematic analysis instead of mere introspection (which is known to be flawed). (3) It may especially be useful for groups of doctors working towards a consensus policy. The material collected in the Drug Education Project was used to compare the decision making policies of GPs in different countries. Regarding treatment of urinary tract infections, it was found that although treatment decisions varied widely between countries, there were remarkable similarities in policies predicting non-optimal decisions in the different countries (Hummerts-Pradier E et al. GPs' treatment of uncomplicated urinary tract infections - a clinical judgement analysis in four European countries. Family Practice 1999;16:605-7.) A similar analysis is now conducted regarding the treatment of asthma.

Petra Denig: p.denig@med.rug.nl

Ecological Rationality

Mandeep K. Dhama (Department of Psychology, University of Victoria, Canada)

Clare Harries and I organized a symposium with a distinctive Brunswikian theme that was presented at the 18th Conference on Subjective Probability, Utility, and Decision Making, European Association for Decision Making, August, 2001, Amsterdam, The Netherlands. Our main aims were to increase the visibility of Brunswik related research at such international meetings, convince some new researchers that they were working on issues that have direct implications for neo-Brunswikian thinking, and stimulate discussion on the concept of ecological rationality. The details of the symposium are in Box 1.

Mandeep K. Dhama: mkdhami@uvic.ca

Testing the dual-mode model of cooperation

Timothy C. Earle, (Western Washington University, USA)

In last year's newsletter, I outlined a dual-mode model that provided an account of how trust (based on shared values) and confidence (based on past performance) interact to produce cooperation. This year, in collaboration with Michael Siegrist and Heinz Gutscher of the University of Zurich, I have begun work on a two-year project designed to test our model. We are only in the initial stages of data collection now, so I will withhold any further comments until next year's newsletter when I will be able to describe the full results of our study.

Tim Earle: timearle@cc.wvu.edu

Box 1. Ecological Rationality in Learning and Decision Mechanisms: SPUDM 18 Symposium

Organisers: Mandeep K. Dhami and Clare Harries

Egon Brunswik (1952) introduced the idea of a Darwinian style adaptation of cognitive processes to the environments in which they function. This idea has been successfully applied to multiple-cue probability learning and the formulation of decision mechanisms. Consequently, the notion that people are ecologically, rather than axiomatically, rational, has become popular. This symposium includes researchers who study the adaptive nature of cognitive processes, namely learning mechanisms in categorical and social environments, and decision strategies in categorical and choice situations. The papers also demonstrate the methods by which ecological rationality may be investigated. The discussants will evaluate and integrate this research.

Discussants: Ulrich Hoffrage and Larry Fiddick.

Judgment in hierarchical learning: conflicting adaptations to the statistical environment: David A. Lagnado and David R. Shanks.

Research in an associative learning paradigm suggests that after exposure to a structured learning environment people give judgments more closely related to predictiveness than to normative probability. This is because their learning mechanisms are attuned to statistical contingencies in their environment, and they use these learned associations as a basis for subsequent probability judgments. Using a medical diagnosis task, we introduced a simple hierarchical structure into this paradigm, setting up a conflict between predictiveness and coherence. Under a probability format participants tended to violate coherence and make ratings in line with predictiveness, and under a frequency format they were more normative. These results are difficult to explain within a unitary model of inference, whether associative or frequency-based.

Ecological constraints on the development of social conventions: Neil Bearden

An assumption underlying evolutionary game theory is that each organism in a population interacts with every other with equal probability. However, this assumption is not met in a natural ecology where organisms are more likely to interact with others closer to them. Experiments investigated geographically constrained, boundedly rational, simple learning organisms that use cues (i.e., past interactions) and a Q-learning mechanism in order to coordinate their behaviors. Through local coordination, complex self-organizing behavior emerges at the global level, demonstrating that a large number of locally rational but globally inefficient conventions can develop under realistic ecological constraints. These observations are not predicted by standard evolutionary game theory, which makes some ecologically implausible assumptions.

Rules and exemplars in human judgment: Peter Juslin

Recent theories of categorization postulate that people simultaneously acquire representations at multiple levels, which compete to control specific responses. When interacting with an environment people acquire knowledge of exemplars and abstract rule-based knowledge of cue-criterion relations (i.e., cue validities). I contrast these systems in terms of their speed and frugality". The exemplar system attains flexibility by storing large amounts of knowledge and postponing all computation until the time of judgment (i.e., a lazy algorithm), whereas the rule-based system compiles knowledge into special-purpose abstractions, thus requiring extensive precomputed knowledge to attain flexibility. I review data from our experiments that aim to ascertain which environments promote either system and what consequences this has for the properties of the judgments.

The effect of interattribute correlations on decision strategies are attribute based or option based: Barbara Fasolo, Gary H. McClelland and Katharine Lange.

In choices among different options, the relationship between the attributes that describe these options determines how we search for information and make a decision. Using a web-based information display board, we found that decision makers responded to positive inter-attribute correlations by using simple attribute-based information search and decision strategies, and to negative inter-attribute correlations by adopting more effortful option-based strategies. Thus, when trade-offs between negatively correlated attributes need to be made, decision makers overcome the inherent difficulty and adopt effortful option-based strategies rather than more frugal attribute-based ones. This reinforces an optimistic view of decision makers able to flexibly adapt search and decision strategies according to the structure of their decision environment.

Personality Judgment and the Realistic Accuracy Model

David Funder and Colleagues, (University of California, Riverside, USA)

Our lab is continuing a stage of intensive data-gathering in pursuit of the factors that moderate accuracy of personality judgment. We are in the process of observing 180 target participants as they interact in 3-person groups, and gathering their self and mutual judgments of personality. We are also using numerous other measures and including a clinical interview and judgment of each participant. The goal is to ascertain the acquaintanceship circumstances and other factors that affect accuracy. The theoretical model underlying this work is the Realistic Accuracy Model (RAM), an adaptation of the Brunswik lens model that views accuracy personality judgment (a.k.a. achievement) as a function of the availability, detection, and utilization of relevant behavioral cues. A recent paper from our lab (Funder, Furr & Colvin, *Journal of Personality*, 2000) describes the technique we developed for the quantitative coding of observations of social behavior. Other work is in preparation.

David Funder: funder@citrus.ucr.edu

Referral Prioritisation in Occupational Therapy and Guideline Use in General Practice

Ken Gilhooly (Brunel University, UK),
Priscilla Harries (Brunel University, UK) and
Liz Smith (Aberdeen University, UK)

KG has been supervising two PhD projects over the past 2 years which apply the lens model approach in two areas of practical importance viz., occupational therapy (PH) and prescribing for depression (LS). We briefly summarise progress so far in these studies.

Occupational therapy

Judgement analysis has been used to capture the referral prioritisation policies of 40 experienced occupational therapists throughout England, Scotland and Wales. The policies are necessary in order to teach undergraduate occupational therapists this vital clinical skill. If an undergraduate occupational therapist chooses to work in a community mental health team they will be overwhelmed with referrals and are unlikely to have an occupational therapy colleague for advice. The novice must have some sense as to which clients require treatment from the occupational therapist. The experienced clinician's policies identify that these judgements of priority have to be based on

such issues as degree of occupational dysfunction, diagnosis, level of risk, and support available. Previous clinical reasoning research in occupational therapy has tended to use ethnographic and process tracing approaches. These have lacked the ability to access the intuitive reasoning of experienced clinicians. Cluster analysis has now been used to identify demographic trends in data. The cluster analysis identified that part-time occupational therapists with greater years of experience were more likely to have a specialist OT role as opposed to a generic role. Whereas less experienced, full time OTs were more likely to have a greater generic role than specialist OT role. Whether undergraduates should be taught the policies of the more experienced clinicians or whether they should be taught the role of those closer to their own potential early career is a current dilemma. The next phase will involve training undergraduate students to use the policies. This is necessary to examine the potential for learning.

Prescribing

The lens model was used to analyse individual treatment policies of 40 General Practitioners (GPs) who were asked to indicate whether they would prescribe anti-depressants for cases varying on 8 cues (e.g., impaired concentration, suicidal thoughts, sleep disorder). Fits to the linear model were good and suggested that most GPs were not following the official non-linear guidelines for prescribing. Also, GPs prescribed at a higher rate than expected on the basis of guidelines. We have also begun to assess the fit of these data to fast-and-frugal models and that approach also appears promising.

Ken Gilhooly: ken.gilhooly@brunel.ac.uk
Priscilla Harries: priscilla.harries@brunel.ac.uk
Liz Smith: mes@hsru.aberdeen.ac.uk

Patient Education

Robert Hamm (Oklahoma, USA)

We are conducting a large patient-education intervention study (315 men, 2-3 hours of questionnaires each) on the topic of prostate cancer screening. Collaborators are our research assistants Manoj Jain, MD, Hesham Aboshady, MD, Renee Patrick (new psych grad student at Ohio State University), David Bard (psych grad student at the University of Oklahoma), Stephanie McFall, PhD (Department of Health Promotion Sciences, University of Albany), Bob Volk, PhD (Department of Family and Community Medicine, Baylor, Houston), Scott Cantor, PhD (Univ of Texas-M.D. Anderson Cancer Center, Houston), Dewey Scheid, MD, Jim Mold, MD, Frank Lawler, MD (my department), and Don Elmajian, MD (Dept of Urology, Shreveport, LA).

Printed educational materials, in an extended balance sheet format, were developed for presenting information to not-recently-screened men concerning prostate cancer (PC) screening. Information about PC prevalence, natural history, PSA and DRE screening accuracy, treatment (radical prostatectomy) efficacy, and treatment side effects was presented. Probabilities of pertinent outcomes were presented graphically, using faces to represent the number of men experiencing each outcome, out of 1000 men screened. Summary pages compared the advantages and disadvantages in side-by-side columns, for screening and for treatment. The numbers and displays were based on a Markov model decision analysis, which was done separately for each ethnic group (white, African American) and each age group (50s, and 60s; and 40s for African Americans).

315 men were randomized to receive the appropriate balance sheet or else to read an NCI pamphlet. Another randomized factor in the design was exposure to more information about the possible outcomes, in the course of a utility assessment procedure. Before and after exposure to the educational intervention, the men completed questionnaires on: intention to get PC screening, knowledge about PC, health beliefs regarding PC screening (benefits, barriers, susceptibility, severity), pros and cons of PC screening and treatment, probabilities of various PC screening and treatment outcomes, and decisional conflict regarding screening.

A quasi Lens Model analysis, with multiple judgments, could be constructed along the following lines.

The criteria could be these several factors:

- a) the objective risk of getting prostate cancer,
- b) the objective risk of getting metastasized PC if you were to get localized PC
- c) the change in objective risk of metastasized PC if you were to treat localized PC
- d) the objective risk of side effects (incontinence, impotence) if treated
- e) the suffering if one were to get metastasized PC, according to report of patients.
- f) the disutility of the side effects, according to report of patients.

The cues would be the subject's risks, displayed to him, and descriptions of the PC outcome state and side effects.

The judgments would be the subject's judgments of the risk factors,

- a) the subjective risk of getting prostate cancer,
- b) the subjective risk of getting metastasized PC if you were to get localized PC
- c) the change in subjective risk of metastasized PC if you were to treat localized PC
- d) the objective risk of side effects (incontinence, impotence) if treated as well as the subjective evaluations of the possible outcome states
- e) the suffering if one were to get metastasized PC.
- f) the disutility of the side effects.

The environmental organizing principle for the probabilities could be the dependency of the risks of cancer upon the cues.

The subjective organizing principle for the subjective probabilities could be the dependency of the judged risks of cancer upon the cues.

Additionally, the objective data could be combined using normative/prescriptive decision analysis into two recommendations: whether to treat if you were to find localized cancer, and whether to screen now.

Correspondingly, we also have subjective statements of intent: whether the subject would get treatment if he had prostate cancer, and whether he intends to get screened for prostate cancer.

There is a problem for the models of the probabilities. The environmental criterion and the environmental model would be based on the same data using essentially the same formula. (This is a problem we faced in modeling "highway traffic carrying capacity" judgments in Hammond, Hamm, Grassia, and Pearson, 1997. We could split it into linear and nonlinear portions. The formula involves logistic regression, while the lens modeling would conventionally use linear regression.)

The most drastic deviation from a Brunswikian design is that we would only have one judgment for each S - for himself. The cues would differ for each subject. So fitting the subjective judgments would be across people, not within people.

The analysis sketched above is not the originally intended analysis, but it is an interesting possibility. Suggestions (or dissuasions) will be attended carefully.

Rob Hamm: robert-hamm@ouhsc.edu

Environmental Structures, Cognitive Structures and the Information Age.

Ken Hammond (Boulder, Colorado, USA)

In this, my 15th year of retirement, and over a half century after my student days in Berkeley, I have witnessed with great pleasure the appearance of "The

Essential Brunswik: Beginnings, Explications, Applications”, which is due largely to the efforts of Tom Stewart. This is essentially the book that Edward Tolman wanted to see produced in 1955 to honor his colleague, but of course, in retrospect, it was better to wait so that we could add the “Applications”, (most of which Brunswik would not have dreamed of) and, in addition, the comments of many researchers (most of whom were not even born at that time). There are few iconoclasts, such as Brunswik was, whose ideas, anathema to the establishment of his time, have been as successful. After “The Essential Brunswik” was launched I began the fourth book of my retirement days. (See? retirement isn’t so bad!) The tentative title is: “Human Judgment in the Information Age: Getting better -- or getting worse?” Don’t let the title fool you. I am just like everyone else; I just keep writing the same thing over and over. But new twists are educational. And I think I can say with confidence that I have finally figured out just what my theoretical approach consists of. And it will be in this book. I have now got the ideas of coherence/correspondence and intuition/analysis straightened out and placed in relation to one another. I have put them to work in structuring the field of J/DM, which, I believe, badly needed it. I will present some of these ideas at the meeting in Orlando for your criticism. I badly need a title for this theory, since it encompasses Cognitive Continuum Theory. There will be a prize for the winning suggestion.

Kenneth Hammond: krahammond@earthlink.net

Strategic and Medical Decision Making: What, When and How

Clare Harries (Leeds University Business School, UK)

Now based in the Centre for Decision Research at Leeds University Business School, West Yorkshire, England, I am pursuing a line of research into the environmental and cognitive factors affecting organisational strategic decision-making. The focus of this research project is to identify which factors, associated with which strategies are most useful for which types of decision-making environment. I have also started new projects on Medical Decision Making with colleagues from the medical school. One project focuses on the role of the environment in guideline formation and normative modelling. Another focuses on means of applying the results of randomised control trials in medical decision making. One key area of investigation will be the role of generalisation across environments.

Clinical Judgment Analysis update. (work with Damien Forrest, Nigel Harvey and colleagues at

University College London, UK) In a ridiculously complex design we’re examining how physicians look for information, order tests, and manage patients who present with chest pain. We compare individual primary and secondary care physicians’ treatment of the same set of patients. We focus on how they select and use age-related information. We analyse this in relation to the characteristics of their practice. In a prequel to this study we used cognitive mapping to model the same physicians’ explicit understanding of the links between the different factors associated with coronary heart disease. Specialist groups differed mainly in terms of their perception of the role of tests. However, a subgroup of General Practitioners had cognitive maps similar to those of Cardiologists.

There will be a one-day meeting on Clinical Judgment Analysis, in Leeds mid-April 2002. Contact me for more information.

Clare Harries: ch@lubs.leeds.ac.uk

Brunswikian Research at the University of Connecticut

Jim Holzworth (Storrs, Connecticut, USA)

Research in the Brunswikian tradition continues at the University of Connecticut.

Lisa Kath completed her Master’s thesis concerning judgments of sexual harassment federal court cases. A manuscript has just been submitted to *Law and Human Behavior* for a special issue devoted to psychology and civil litigation. U.S. federal court rulings and layperson judgments of severity and pervasiveness of 50 real sexual harassment court cases were investigated using judgment analysis (policy capturing). Correlations (agreement indices) between court rulings and layperson judgments were significant for judgments of 50 out of 53 (94%) jury-eligible Americans. This supports the hypothesis that there is agreement between legal and psychological definitions of severity and pervasiveness. To examine the nature of this agreement, judgments were regressed onto five cues (status of harasser, level of coercion, level of physical contact, number of incidents, and vulgarity). These cues accounted for statistically significant policy models for 42 (79%) laypersons, and a statistically significant policy model for court rulings. No gender differences in judgments or agreement were found. We are now investigating factors that influence judgments of employer liability for alleged incidents.

Kris Canali and I are interested in comparing methods of judgment analysis. Gary McClelland (1999) hypothesized that using fewer cases with more extreme cue values (efficient research designs) may be more statistically efficient, and therefore more expeditious, than traditional representative designs in judgment analysis. In our first study, 64

undergraduates judged attractiveness of either 30 efficient or 40 representative descriptions of actual roommates based on eight cues, and then made 30 additional attractiveness judgments about representative cases for cross-validation. The efficient research design accounted for significantly more judgment variance in the judgment analysis stage (produced larger squared multiple correlations), and there were no negative consequences on cross-validated squared multiple correlations.

Jim Holzworth: holz@uconnvm.uconn.edu

Modeling Fault Diagnosis in a Dynamic Process Control Task using a Multivariate Lens Model

Pratik Jha and Ann M. Bisantz (Buffalo, The State University of New York, USA)

This work has been supported by NSF Grant #IIS9984079 and was completed as part of the first author's M.S. thesis. The description below summarizes work that was presented at the 2001 Human Factors and Ergonomics Society Conference.

Models of dynamic or naturalistic decision making emphasize the cyclical nature of decision making in the real world. Recognition or judgments of the state of the environment are a critical component of these models. Because the Lens Model can be used to describe judgments, including those of system state, and it provides a means to quantify aspects of judgment performance, it has potential value in modeling the components of dynamic decision making that involve recognition of the situation. While there has been extensive application of the univariate Lens Model, which relates single judgments to single environmental criterion, many judgments of interest are multivariate, or categorical in nature. For instance, one may want to determine how a situation rates on each of several variables, or judge which of several situations are occurring. To address such situations, a multivariate extension of the Lens Model has been developed (Castellan, 1972; Cooksey and Freebody, 1985). Analogous parameters to those identified for the univariate Lens Model above can be computed for the multivariate Lens Model, using canonical correlation procedures. However, while the univariate Lens Model has been applied in numerous contexts (see Cooksey, 1996) such as medical decision making and social policy judgments, the multivariate Lens Model has few demonstrated applications (see Cooksey and Freebody, 1986 for an example in a social judgment domain), and none known for a dynamic situation such as fault diagnosis in process control. Therefore, the motivation of the present study was to

investigate the effectiveness and applicability of a multivariate Lens Model to human judgment in a complex, dynamic task: fault diagnosis in process control. We also investigated sensitivity of this modeling technique to capture variations in the participant's performance. Results of such modeling could be used to evaluate the effectiveness of displays, and through an assessment of operator performance, provide information regarding training needs. Our approach to the research was to conduct both a sensitivity analysis to determine model responses to simulated fault diagnosis data, and to collect and model experimental data (Jha, 2001). To support the fault diagnosis tasks, we utilized an existing, dynamic process control simulation called DURESS II (Vicente, 1999; Christofferson, Hunter, and Vicente, 1996). Twenty-four faults, at two levels of change from steady state, were created, for a total of forty-eight faults. Faults could be categorized as one of nine categories reflecting the component that failed, and the type of fault. Cues were identified and measured based on information available through each of two graphical interfaces to the simulation: one interface provided 4 cues, while the second provided 20 cues. The task was to watch the simulation and categorize the fault based on information available from the interface. For the sensitivity analysis, three sets of simulated judgments were created: one with 25% random judgment errors, one with 50% random judgment errors, and one with 75% random judgment errors. Cue values were captured one minute after a fault occurred, and multivariate lens model parameters were calculated for each of these judgment sets. Sensitivity analysis results indicated that model parameters were sensitive to simulated performance differences: for instance, the measure analogous to r_A changed from .94 to .85 to .69 as errors increased. For the experimental task, 16 participants performed 48 diagnosis trials in one of the two interface conditions. Unexpectedly, while performance ranged from 89% to 35% correct, r_A ranged between .9 and .99. This result appeared to be due to the nature of the canonical correlation procedure. Canonical correlation expresses the strengths of relationships between subsets of variables within two sets of interest (in this case, judgments and actual faults). Even for poor performing participants, there was at least one strongly related subset (i.e., there was at least one significant root) reflecting the fact that they were making some judgments correctly. Further inspection of the canonical outputs showed that better performing participants had more sets of more strongly related variables (a greater number of significant canonical roots) than poorer performing participants. Inspection of matrices of canonical weights allowed an interpretation of diagnosis performance in terms of which judgment categories were associated with which actual fault categories. These interpretations were

consistent with the patterns of judgments and faults. In contrast, the data sets created for the sensitivity analysis exhibited random patterns of errors and (for the poorest performing case) no strong judgment-fault pairing. Thus, for poorer performing simulated data, there were no strong associations and no significant roots.

In summary, the study demonstrated the potential applicability of the multivariate lens model to a reasonably complex, dynamic environment. Sensitivity analyses showed that Multivariate Lens Model parameters did vary based on level of judgment performance, while experimental results indicated less sensitivity of model parameters to poor performance because some faults were diagnosed correctly. Further inspection of outputs of canonical correlation procedures allowed identification of these systematic strategies.

Ann Bisantz: bisantz@eng.buffalo.edu

Judgment Analysis of Quality of Life Assessments

Dick Joyce, (Dublin, Ireland)

Dick Joyce, with colleagues in (1) Bern, (2) Zürich and (3) Dublin, is striving to complete and even (with editors' and referees' indulgence) to publish his last three swansongs. All, of course, have to do with the application of JA to the assessment of individual quality of life; (1) that of alcoholics treated with disulfiram (Antabuse) and social support, (2) of HIV+ people given Chinese herbal medicine as an adjunctive treatment to conventional therapy, and (3) a history of the Schedule for the Evaluation of Individual Quality of Life (SEIQoL).

Dick Joyce dick.joyce@ireland.com

Cognitive Feedback and Musical Expressivity: Teaching the Thing that Cannot be Taught

Patrik N. Juslin, (Uppsala University, Sweden)

This year I had the pleasure of launching a new research project with distinctly Brunswikian flavors. The project is called Feedback-learning of Musical Expressivity (Feel-ME), and the purpose of the project is to (a) define the nature of expressivity in musical performance in computational models, and (b) develop new methods for teaching expressive skills based on recent advances in musical science, psychology, technology, and acoustics. Many studies suggest that expressive aspects of performance are neglected in music education, and that traditional methods for teaching expressivity rarely provide specific feedback to a performer. Part of the problem is that communication of emotions

involves tacit knowledge about complex relationships among performer intentions, acoustic cues in music performances, and listener judgments. Knowledge about such relationships is difficult to convey from teachers to students. Hence, many teachers and students regard expressive skills as "the thing that cannot be taught". In the Feel-ME project, a novel approach to learning musical expressivity, based on Brunswik's Lens Model and Hammond's concept of Cognitive Feedback, is developed and implemented in software. The project involves an inter-disciplinary collaboration among psychologists, technicians, teachers, and musicians. Members of the group are currently involved in developing computer algorithms for automatic extraction of acoustic cues in music performances. The results from these analyses are modeled partly on the basis of judgment analysis procedures. The work has been successful so far, so we hope to test early prototypes of the software on music students next year. We also hope that the software will allow us to study the actual process of learning cue-relationships in musical communication, by storing each learning stage in terms of various models and statistics. This year also saw the publication of the first multi-author book on "music and emotion", co-edited by myself and John Sloboda. The role of emotion in music has been a subject of considerable debate ever since Antiquity. Yet, emotional aspects of music have received less attention than cognitive aspects in music-psychological research. A new volume in the Series of Affective Sciences, Music and Emotion brings together leading researchers interested in both these topics to present the first integrative review of this subject. The volume features various multi-disciplinary perspectives, taking on board views from philosophy, psychology, musicology, biology, anthropology, sociology, and therapy. The Brunswikian perspective is - needless to say - most prominently featured in my own chapter ...

Selected publications:

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1797-1813.

Juslin, P. N., & Sloboda, J. A. (Eds.) (2001). *Music and emotion: Theory and research*. New York: Oxford University Press.

Juslin, P. N., Friberg, A., & Bresin, R. (in press). Toward a computational model of expression in music performance: The GERM model. *Musicae Scientiae*.

Juslin, P. N., & Persson, R. S. (in press). Emotional communication. In R. Parncutt, & G. E. McPherson (Eds.), *The science and psychology of music*

performance. *Creative strategies for teaching and learning*. New York: Oxford University Press.

Patrik Juslin: patrik.juslin@psyk.uu.se

Process and Representation in Multiple-Cue Probability Learning

Peter Juslin, (Umeå, Sweden)

In a series of recent experiments, we have returned to the classical Brunswikian domain of multiple-cue probability learning. The aim of these studies is to connect this rich body of data with more recent theorizing in cognitive science, and research on category learning, in particular. In contrast to most research on multiple-cue probability learning that has mainly been concerned with descriptive analysis—for example, in terms of the lens-model—research on categorization has produced an extensive list of detailed models of the cognitive processes and representations (although, arguably, at the cost of ignoring environmental issues). In our studies, we attempt to relate multiple-cue judgment to both exemplar models from categorization research and rule-based models that presume mental integration of cue-criterion relations at the time of judgment. The possibility of exemplar retrieval is illustrated by a physician that makes diagnosis by retrieving similar previous patients with known diagnosis. The possibility of rule-based, mental cue integration is illustrated by a physician that considers multiple specific symptoms (cues) that are integrated in the light of known cue-criterion relations to make an overall diagnosis. In the experiments, we try to ascertain the relative importance of each kind of process as a function of environmental parameters, and how this relates to, for example, judgment performance and lens-model indices. We hope that this may contribute to a more detailed and refined understanding of such key-concepts in Brunswikian psychology as “quasi-rationality”, “intuition” and “analysis”.

Peter Juslin: peter.juslin@psy.umu.se

The Continuing Search for Configurality

Pieter Koele (Amsterdam, The Netherlands)

In structural modelling research the linear-additive model seems to be unbeatable as the best description of multi-attribute (multi-criteria, multiple cue, whatever you prefer) judgement/evaluation strategies, this in spite of the fact that quite a few professionals claim that their judgement strategies are far more complicated than this model suggests.

A so-called configural strategy is one of these more complex strategies. It is a strategy in which the way a particular attribute is used depends on the value of another attribute. Conjunctive and disjunctive

strategies are well-known examples of such strategies. Attempts to model conjunctive and disjunctive strategies have been undertaken by, among others, Einhorn, Brannick and Ganzach. In general, the fit of their models is hardly any better than the fit of the linear-additive model.

In a research project I am about to start I will further specify configurality, by distinguishing between strategies in which the value of a particular attribute influences the weight of another attribute (like in the conjunctive and disjunctive strategies), and strategies in which the form of the relation between the other attribute and the judgement is influenced. I will first investigate whether judges are capable of handling such strategies, when they are explicitly instructed to do so. Individual difference will be taken into account. Next I will investigate in an MCPL context whether and under which conditions judges are able to detect the configurality that is present in an ecological system.

Pieter Koele: ml_koele@macmail.psy.uva.nl

Representative Design Revisited

Gary McClelland and Barbara Fasolo (University of Colorado at Boulder, USA)

As Jeryl Mumpower noted in his summary, I sometimes come to the Brunswik meetings to play devil's advocate--and I'm not just pretending. But as Ken Hammond points out to me, I haven't read enough Brunswik to know what he really advocated about representative design. Perhaps *The Essential Brunswik* will improve my education. I instead have been making my inferences about representative design from the research behavior of those who call themselves Brunswikian. I infer that an important feature of representative design is the faithful representation of the statistical properties (means, variances, and especially correlations) of the environment in the judgment tasks we use (although this is talked about more often than it is actually done). It seems to me that there are two reasons that representative design might be important: (a) to be able to assess the judge's achievement or the correspondence between judgments and the actual values and (b) to not distort the cognitive processes or "policy" the judge would normally use. As I've argued at previous meetings, I believe (a) is an undesirable reason because it leads to statistically inefficient designs that can only "reveal" simple judgment models and, relatedly, it does not allow us to examine a judge's correspondence for those cases that might be most important (e.g., how a physician or a pilot handles a rare, but dangerous, case). Also, inefficient representative designs require so many judgments as to try the patience of all but the most dedicated of the judges we study. Furthermore, this notion has its roots in statistical ideas that are now outdated. We now

know how to estimate the parameters of a statistical model using an efficient design, and then extrapolate to estimate how that model performs in an environment with any other specified design. From an earlier report, it appears that Jim Holsworth and Kris Canali have demonstrated this empirically. Hence, from a statistician's viewpoint, (a) is now a non-issue. But (b) is potentially a more serious problem if judges adapt (as Payne and colleagues have long argued) their judgment and decision strategies as a consequence of the statistical properties of the task. On this issue, Barbara Fasolo and I will be coming to this year's Brunswik meetings as the opposite of the devil's advocates (I'll let Jeryl or someone else suggest what that is). In studies of web-based consumer decision aids, we've found that consumers rapidly switch their information processing strategies as a function of the average correlation among product attributes and that various decision aids are differentially effective and satisfying to consumers depending on the correlational structure. In research we reported in last year's newsletter, we were frankly stunned at how quickly consumers switched from attribute-based strategies (consistent with non-compensatory processes like elimination-by-aspects) to option-based strategies (consistent with compensatory strategies such as weighted averaging) as we changed the correlational structure from positive (a "friendly" choice environment) to negative (an "unfriendly" environment requiring tradeoffs). More recently we've investigated the effect of attribute correlation on the effectiveness of EBA and rating (or "weight-and-add") decision tools (dare I say these tools are "representative" of those found on the web and used by millions of consumers). With positive correlations, we found both site designs to be helpful and liked. However, with negative correlations: (a) choices were reached more slowly (with more clicks), and with less confidence and satisfaction than with positive correlation, regardless of the site design, and (b) negative perceptions and the need to use more features of the decision site were greater on the site with a non-compensatory attribute focus. This confirms that with negative correlations users detect the conflict among attributes, and have a preference for a more compensatory/option based choice process. Making tradeoffs is hard though, and users are happy if they can let the decision tool do the heavy lifting, as it is possible on the compensatory decision site. Our next step in this research program, in conjunction with the Center for Adaptive Behavior and Cognition at the Max-Planck Institute for Human Development in Berlin, is to develop and test a third "fast and frugal" site design that combines the advantages of both designs, making the cognitive and motivational demands more similar to those required by an

attribute focus while also guaranteeing that the choice made is the best possible if all attributes were considered as it would follow from a compensatory process.

Barbara Fasolo: barbara.fasolo@colorado.edu

Correspondence, Coherence and the Cognitive Continuum

Kathleen Mosier (San Francisco State University, USA)

This year, I have been striving - with much help from Ken Hammond - to re-orient my thinking about decision making in automated aircraft in terms of coherence/correspondence, and the CCT. An outline follows:

The shift from active control to systems monitoring in the automated cockpit has profoundly changed the type of cognitive activity required of pilots. Most importantly, in terms of theoretical implications, the automated cockpit brings cues that were in the outside environment into the cockpit, and displays them as highly reliable and accurate information rather than probabilistic cues. This changes the goal of pilot cognition from correspondence, or empirical accuracy in using probabilistic cues for diagnosis, judgment, and prediction, to coherence, or rationality and consistency in diagnostic and judgment processes (Hammond, 1996; 2000). In contrast to earlier pilots, glass cockpit pilots can spend relatively little of their time looking out the window, and most to all of it focused on information inside the cockpit. The data that they utilize to fly can, in most cases, be found on cockpit display panels and CRTs. These data are qualitatively different from the cues used in correspondence judgments. They are data, rather than cues - that is, they are precise, reliable indicators of whatever they are designed to represent. The cockpit is a deterministic, rather than a probabilistic environment, in that the uncertainty has, for most practical purposes, been engineered out of it through high system reliability. In the automated cockpit, then, the priority for correspondence in cognitive processing has been replaced by a demand for coherence. This shift in cognitive goals means that we need to re-examine cognition in the automated cockpit to determine what is required to achieve, maintain, recover coherence in the cockpit, and whether or not these processes are supported by current displays of information.

Intuition and Analysis.

The goals of correspondence and coherence can be achieved by cognitive tactics ranging on a continuum from intuition to analysis (e.g., Hammond, 1996; Hammond, Hamm, Grassia, & Pearson, 1997). Within the automated cockpit, the design and display of most automated systems elicits intuitive cognition. This

design philosophy seems to be consistent with the goals of workload reduction and information consolidation. However, current displays may in fact be leading pilots astray by fostering the assumption that cockpit data can be managed in an intuitive fashion. Although pilots can intuitively infer coherence among cockpit indicators much of the time if things are operating smoothly, repairing - and often detecting - disruptions to coherence demands a shift toward analysis. Many errors and anomalies, such as being in the incorrect flight mode, can only be detected via analysis. Additionally, the complex nature of the automated cockpit requires that disruptions to coherence be resolved via analytical means. Data in displays must be compared with expected data to detect discrepancies, and, if they exist, analysis is required to resolve them before they translate into unexpected or undesired aircraft behaviors. Displays in the cockpit, then, should not only support intuitive processes, such as the quick detection of some out-of-parameter states, but must also provide the information necessary for analysis. If the pilot is expected to maintain coherence in the cockpit, he or she must be able to develop accurate mental models of system functioning. In order to track system status and resolve anomalies, the electronic world must support analysis of current states and resolution of discrepancies.

Kathleen L. Mosier: kmosier@sfsu.edu

Intellectual Activities of a Brunswickian Administrator

Jeryl Mumpower, (Albany, New York, USA)

As a condition of employment, upon becoming an Associate Provost and Graduate Dean, I (Jeryl Mumpower) had to promise that I would not initiate any intellectual work. I found a slight loophole, however. My administrative pledge said nothing about finishing up work that I had commenced years and years ago.

This past Spring, the *Journal of Policy Analysis and Management* accepted a paper by Radhika Nath, Tom Stewart, and myself, entitled, "Affirmative Action, Duality of Error, and the Consequences of Mispredicting the Academic Performance of African-American College Applicants." Some of you were at the 1999 Boulder meeting, where I presented some of this research, but my earliest files date back to 1997. This research takes a distinctly Brunswickian perspective, as I've indicated in previous updates. To understand the full implications of different affirmative action policies, one must look to the task. Four factors must be considered simultaneously: selection rate, or percentage admitted; base rate, or percentage of those applying who could do the work if admitted; predictive accuracy, or degree of

correspondence between predictions of performance and actual performance; and the costs associated with false positive and false negative errors, as well as the benefits associated with true positive and true negative diagnoses.

In August, Gary McClelland and I completed a distinctly Brunswickian paper entitled "Measurement Error, Skewness, and Risk Analysis: Coping with the Long Tail of the Distribution." We are pleased that this paper - for which my earliest drafts say in faded, barely legible type "July 1990" - has been accepted for publication by *Risk Analysis*. Yes, that's the same Gary McClelland who sometimes comes to our meetings to play devil's advocate (and he's not just pretending.) But on this issue we are in good Brunswickian accord. The paper analyzes the task environment that faces the judge in risk analyses and decision lotteries. If the analysis requires multiplication of variables that contain random error, then resulting estimates of expected value, expected utility, or joint probability may be skewed, sometimes highly so. Taking a Brunswickian approach to the analysis led Gary and me to some interesting and sometimes surprising conclusions:

- Joint probability estimates based on the analysis of multi-stage chance trees are more likely than not to be below the true probability of adverse events, but will sometimes substantially overestimate them.
- In contexts such as insurance markets for environmental risks, skewed distributions of risk estimates amplify the "winner's curse" so that the estimated risk premium for low-probability events are likely to be lower than the normative value.
- Skewed estimates of expected value may result in risk-neutral decision makers exhibiting a tendency to choose a certainty equivalent over a lottery of equal expected value, or vice versa.

When distributions of estimates of expected value are positively skewed, under certain circumstances it will be optimal to choose lotteries with nominal values lower than the value of apparently superior certainty equivalents.

Jeryl Mumpower: JMumpower@uamail.albany.edu

Analyzing the Utility of Utilities and Conceptual Disease Spaces

Radhika Nath (Albany, New York, USA)

I am currently involved in two studies. In one study, I am examining how individual differences affect utility

estimates for different health categories obtained using various utility elicitation instruments. Given the unresolved debate in the area of medical decision making concerning utility estimation, I believe this study is overdue. This study has implications for inclusive decision making, medical decision making and the social policy creation process. I am currently in the process of data analysis and results will be forthcoming.

I am also in the process of wrapping up another study where I undertook a multi-dimensional scaling analysis of several diseases in order to understand individual representation of a conceptual health/disease space.

In the past year I have seen studies with Drs. Mumpower and Stewart and with Dr. Stewart, respectively, come to completion. The former study was a Taylor-Russell analysis of affirmative action in College admission decisions which was accepted for publication at the *Journal for Policy Management*. The latter was a report prepared for the Forecast Systems Laboratory of the National Oceanic and Atmospheric Administration regarding the value of improved forecasts.

Radhika Nath: RN5534@albany.edu

Fast and Frugal: Some of the People, Some of the Time, Some of the Environments?

Ben Newell (University College London, UK)

In a project starting in April of this year David Shanks and I have been conducting a number of experiments aimed at testing the empirical validity of the fast and frugal heuristics approach (e.g. Gigerenzer et al. 1999). Our initial focus has been on the 'take-the-best-heuristic' and the abstract of our recently submitted paper appears below (Box 2). Our findings suggest some people are 'fast and frugal' some of the time with others integrating information from all available cues in the environment. Current and future work aims to investigate factors that might affect the adoption of these 'frugal' or 'weight of evidence' strategies, such as the total number of cues in the environment and the relative costs involved in learning about and acquiring cues. The work is part of the programme of the ESRC Centre for Economic Learning and Social Evolution.

Ben Newell: b.newell@ucl.ac.uk

Box 2. Take the best or look at the rest? Factors affecting 'one-reason' decision-making Poster presentation, JDM 2001

Ben R. Newell and David R. Shanks.

The "Take-the-Best" (TTB) heuristic (e.g., G. Gigerenzer & D.G. Goldstein, 1996) states that when making a choice between two alternatives, people will base their choices on what they perceive as the most valid or 'best' piece of information that discriminates between the alternatives. We report three experiments in which aspects of the experimental environment were manipulated to examine the parameters under which such a strategy operates. Clear evidence of TTB use was detected in all the experiments. However, the experiments also demonstrated that even under conditions contrived to promote the use of TTB low cost of information, prior instruction as to the validity of cues, and a deterministic environment there was a high proportion of behavior that was inconsistent with TTB. Together with the presence of large individual variability in strategy use in the three experiments, these results question the validity of TTB as a psychologically plausible and pervasive model of behavior.

Evaluating Expertise

James Shanteau (Kansas State University, USA)

What is CWS? CWS is the Cochran-Weiss-Shanteau approach to assessing expertise purely from data. The approach is based on the idea that expert judgment involves discrimination – seeing fine gradations among the stimuli – and consistency – evaluating similar stimuli similarly. The approach was inspired by an idea for comparing response instruments suggested by the late statistician William Cochran (1943), and adapted to the domain of expertise by David J. Weiss and James Shanteau.

What do we mean by an "expert"? Applying the term "expert" to a person is a shorthand description of a set of results rather than a characterization of the person. Talent and training may combine to yield a person we label as expert, but it must be kept in mind that the label is a generalization. It is the behavior that is, or is not, expert. With CWS, we measure expertise in a specific setting, with specific stimuli and a specific task. Someone who excels in one context may not excel in others that seem similar.

What is the CWS Index? The CWS index is a numerical value that captures the degree of expertise demonstrated in a set of responses. It consists of the ratio of discrimination and inconsistency.

Discrimination refers to the candidate's differential evaluation of the various stimuli within a set. Consistency refers to the candidate's evaluation of the same stimuli similarly over time; inconsistency is its complement. The ratio will be large when a candidate discriminates effectively, and will be reduced if the candidate is inconsistent.

$$\text{CWS} = \text{Discrimination} / \text{Inconsistency}$$

The rationale for incorporating discrimination and consistency into an index of expertise is that a good measuring instrument, such as a ruler or a thermometer, has these properties. Discrimination and consistency are building blocks of measurement. Similarly, expertise at its core requires the ability to evaluate the stimuli in one's domain. Note that accuracy is not involved in CWS, as we do not assume any knowledge of correct responses.

How are Discrimination and Inconsistency measured? For data obtained using interval or ordinal scales, we usually measure both quantities as variances. Discrimination is the variance among averaged responses to different stimuli, while inconsistency is the variance among responses to the same stimulus, averaged across stimuli. The variance, because it entails squaring deviations, has the property that large differences are accentuated. (It is also plausible to measure discrimination and inconsistency with other dispersion measures, such as standard deviation or mean absolute deviation.)

Because different stimuli are generally expected to elicit different responses, the measures of discrimination and inconsistency are strictly tied to the set of stimuli actually presented. It is not meaningful to compare measures obtained using different stimulus sets. Further, the CWS index is scale independent; a linear transformation of the response scale does not alter the ratio.

Why is the CWS Index set up as a ratio? The ratio formulation reflects the idea that a candidate can trade off the two quantities as dictated by the situation. By widening the range of responses used, one can increase discrimination, but only at the cost of decreasing consistency. Everyone strikes a balance between discrimination and consistency. Performing well in one respect or the other is relatively easy. Someone who can do both at once is behaving expertly.

What is the unit of analysis – Who is “the candidate”? In general, an individual person, a candidate expert, generates a single CWS score for a particular experimental condition. That CWS score may be compared to the score produced by other candidates under identical conditions, or to the CWS score produced by the same candidate under a

different experimental condition. It is also possible for a team to produce a single CWS score, when it is the team's responses to the various stimuli that constitute a data set. In this usage, components of the response from individual team members would not be analyzed separately. The CWS score from one team may be compared to that from another team, or to the CWS score produced by an individual operating alone with the same stimuli.

What kind of research design is needed? In order to be able to distinguish among candidate experts, it is a good idea to present a wide range of stimuli. The researcher may not know how to identify stimuli that span the range; subject matter experts (SMEs) may be useful in selection of stimuli (although, in general, we do not wish to assume expertise, we acknowledge that SMEs do exist and can be helpful). The wider the range of stimuli, the more discrimination it is possible to exhibit. In some cases, expertise may show itself only when rare problems come along. It is crucial that at least some stimuli be presented more than once. This repetition is necessary to provide an estimate of consistency. If it is not practical to present the entire set repeatedly, it is a good idea to select values from across the stimulus range to be presented repeatedly.

There is a danger that amount of inconsistency depends upon stimulus magnitude (à la Weber's law). If it is feasible, we recommend complete replication (presenting the entire set of stimuli more than once); the more replications, the more reliable the estimate. Whether responses in the research setting can be sufficiently isolated to approximate independence is a standard concern for researchers.

It is imperative that the same set of stimuli be presented to all candidates. Varying the order of presentation of those stimuli, perhaps by employing independent randomization, across candidates may be acceptable, if order doesn't have an impact of its own.

What are the constraints on the stimuli to be evaluated? Stimuli need to be presented identically to all candidates. For ephemeral stimuli, that may present a practical problem. Video recording is a valuable tool, although some information may be lost relative to live presentation. For stimuli that are consumed during the task, sufficient quantities need to be on hand and they must not decay over time. Stimuli also need to be presented more than once to an individual. If the stimuli are memorable, the candidate may try to recall rather than render an independent response. The researcher may need to space trials over time or re-label them so as to inhibit recollection.

What sorts of responses are usable? Because our approach to evaluating expertise is quantitative, we require discrete responses. These can be expressed on any of the response scales that experts use. With

numerical ratings, the responses are approximately on an interval scale and can be used to construct variance ratios. Ordinal responses, such as rankings or letter grades, can be used similarly when converted to numbers. Sufficiently dense ordinal data yields considerable power (Weiss, 1986). Even "Yes-No" responses can yield results essentially equivalent to those obtained with continuous scales (Lunney, 1970). Nominal responses generally provide less power; nevertheless, measures of discrimination and inconsistency have been defined.

What statistical procedures are applicable? When CWS estimates of discrimination and inconsistency are variances, there is a statistical comparison that provides significance statements. Schumann and Bradley (1959) developed a procedure for comparing sets of data from two similar experiments that can determine whether one F-ratio is significantly larger than the other. Similar means that each F-ratio is constructed from responses to the same stimuli and therefore, has the same degrees of freedom. The test can be employed either directionally or nondirectionally. The one-tailed (directional) test determines whether the candidate is significantly less capable than a designated expert. The two-tailed (nondirectional) test asks whether there is a significant difference between two judges.

Each judge is considered as a separate "experiment". A computer program (Weiss, 1985) incorporating the Schumann and Bradley procedure and table of critical values is available. The test allows comparison of the expertise exhibited by the various candidates as they judge a particular set of stimulus objects. Pairwise comparisons express how each candidate does with respect to the others. Alternatively, one may compare the candidate's expertise to an established standard.

Example

A recent study by Skånér, Strender, & Bring (1998) illustrates how expertise can be evaluated based on a set of judgments. Twenty-seven Swedish General Practitioners (GPs) judged the probability of heart failure for 45 cases based on real patients; five of the cases were repeated, although the GPs were not informed of that. The case vignettes stated that each patient came to the clinic because of fatigue. There were no additional pathological findings based on further examination. Normal values were provided for hemoglobin, electrolytes, s-creatinine, and TSH. Case-specific information was provided for ten cues: age, gender, history of myocardial infarction, heart X-Ray, and lung X-ray.

"For each vignette, the doctors were asked to assess the probability that the patient suffered from

any degree of heart failure" (Skånér et al., 1998, p. 96). The assessments were made on a graphic scale with "totally unlikely" at one end and "certain" at the other; these were converted into 0-to-100 values. The doctors were instructed "to judge the probability, not the severity, of heart failure" (p. 96).

CWS scores were computed for each of the GP's. The results revealed considerable variation between and within the four GP's. Still, each GP shows a distinctive pattern in terms of discrimination and reliability. CWS scores ranged from 580.20 (= 3365.15/5.80) to .76 (= 330.40/434.00). Some GP's showed high discrimination and low consistency, whereas others showed low discrimination and high consistency. Relatively few doctors, however, had both high discrimination and high consistency.

Based on these results along (with no knowledge of the correct answers), we can gain considerable insight into the judgment strategies and abilities of the GP's. Dr. Skånér and her colleagues in Sweden are now following up this approach.

More information about CWS, along with computational routines and sample datasets, is available at our website: <www.ksu.edu/psych/cws>

James Shanteau: shanteau@ksu.edu

Exploration of Alternative Designs for Judgment Analysis Application in Public Policy Formulation

Junseop Shim (Albany, New York, USA)

My current research concentrates on a dissertation that addresses some methodological issues on judgment analysis, especially on the design of JA study. Drs. Tom Stewart and Elise Weaver have been actively involved in this research. As a doctoral student in the department of public administration and policy, I conducted a literature review on the application of judgment analysis (JA) in public policy area. From this, I found one important constraint that may prevent JA from being widely applied to public policy formulations and other policy areas. To be simple, why does JA require too many cue profiles that burden the judges such as policy stakeholders?

In order to answer this question, my study has two primary purposes. First, it investigates the minimum number of cases required to generate stable model coefficients for judgment policies given the properties of judgment and task. For this, I am conducting a series of statistical computer simulations. I generated a baseline judgment model by mimicking a real JA study. From the baseline model, by varying policy characteristics (R^2 , the number of cues used, function forms) and task conditions (the number of cues in the profile, cue intercorrelations), I am investigating how

such variations affect the stability of parameter estimates, thereby obtaining a generalizable idea about the minimum number of cases. The stability of regression coefficients (beta weights) is measured by the standard error of the coefficient.

Second, it explores a substantive coupling between the representative principle and statistical efficiency in obtaining a given level of stability in designing a judgment analysis study. McClelland's (1999) essay, "Representative and efficient designs" motivated me to investigate a possibility of coupling between those two contradictory principles. For this, I am comparing three different design algorithms (conventional representative design, augmented represented design (ARD), and efficient plausible designs (EPD)) and thereby comparing the relative efficiencies of them. Since, however, ARD and EPD require an appropriate sampling algorithm that allows drawing only extreme or efficient cases from the population of cases in the ecology, I developed four different sampling algorithms, termed classified maximum leverage (CML), Maximum Leverage Distance (MLD), classified standardized efficiency score (CSES), and maximum SES distance (MSESD)"

A Visual Basic for Applications (VBA) program in Excel, called "JASIM" was written to generate and analyze the simulation data. The program consisted of several parts such as cue generation, error generation, judgment generation, regression analysis, the bootstrap, permutation test, and output. As statistical techniques, I employed two resampling algorithms (the bootstrap and permutation test). Therefore, it will be possible to compare the theoretical standard errors of parameter estimates with those calculated from the sampling distribution by both resampling methods, thereby examining discrepancies among them.

I am still conducting simulations. I will be able to report important results to the next newsletter.

Junseop Shim: Js7572@albany.edu

Modeling & Predicting Cardiac Disease in "low risk" Patients with Chest Pain

Tom Tape (Nebraska, USA)

The chest pain research group at University of Nebraska has been working on modeling physicians' decisions whether to admit patients with "low risk" chest pain. Patients who present to the emergency department with acute or recent chest pain are quickly evaluated to determine whether they are having a heart attack. Those without signs of obvious heart attack comprise the group we call "low risk". More than likely, their symptoms are due to a non-cardiac condition, but a non-negligible number

(10-20%) will nevertheless be found to have an acute cardiac condition. For this reason, many of these patients are admitted to the hospital for 24-48 hours to "rule out" an acute cardiac condition. Considerable work has been done by others developing "chest pain protocols" to expedite the inpatient evaluation of such patients. We are interested in studying physicians' decision making regarding which patients they choose to admit. We studied a random sample of patients who were seen for chest pain. We prepared transcripts of their emergency department records. We recorded the actual admitting decision and also had the charts reviewed by two independent physicians who indicated whether they would have admitted the patient (the reviewers were blinded to the actual decisions). We also determined the presence or absence of 44 cues that are potential predictors of acute cardiac disease. Using the cue set, we were able to model both the actual decisions and the decisions of the two reviewers. We compared a number of fast and frugal methods (that we modified to classify the patients) to logistic regression modeling of the admitting decision. The details of our work will be presented at the Brunswick Meeting in Orlando this fall (see the back page of this newsletter – Ed.). The bottom line from the analysis was that none of the fast and frugal methods worked as well as logistic regression (although Dawes rule came close). We theorize that the large number of weak cues favors the regression methodology. We are now interested in testing these methods in other medical decisions, particularly ones with stronger and more dominant cues.

Tom Tape: tgtape@unmc.edu

Brunswikian Accounts of Scientific Thinking

Ryan D. Tweney (Bowling Green State University, USA)

During the past year, I have been actively pursuing a project that doesn't seem, on the surface, to be very related to the interests of this society, but which, I've come to believe, has deep Brunswikian themes.

As part of a long-running project to understand the scientific thinking of Michael Faraday, I've begun an intensive analysis of nearly 700 microscope slide specimens prepared by Faraday during the course of a year's research (in 1856) on the optical properties of gold and other metals. Because these specimens survive, and are accompanied by Faraday's characteristically thorough diary records, we have the opportunity to track the experimental and theoretical activity of a major scientist at a level of detail that I believe is unprecedented. (Examples can be seen at my website <http://personal.bgsu.edu/~tweney>). The "Brunswikian" angle derives from the fact that we can now characterize both his cognitive activity (from the

diary) and the environmental "cues" that he used to structure and guide his thinking. Much of this thinking is visual and, without the specimens, there is simply no way to understand what was driving his activity.

Over the past year, I have concentrated on learning the photomicrographic skills that will allow a modern reader to "see" the specimen archive, and my students and I have begun to replicate some of the chemical preparations used, in order to further understand the "hidden" or "tacit" aspects of Faraday's work. So far, we have managed to (1) use electrical currents to explode pure gold wires, producing deposits on specimen slides that are much like Faraday's, (2) produce colloidal gold suspensions that behave exactly like Faraday's, and (3) have carried out some precipitations of gold in order to replicate his optical examination of precipitates. We have yet to prepare slides of metallic films that use his methods, partly because the substances used are extremely hazardous, but these experiments are in the works. In Brunswik's terms, it could be said that we are presently examining the ecology within which Faraday worked, and we are beginning to see how to incorporate knowledge of this ecological environment in a cognitive account of his thinking during the course of the research. Others, of course, have replicated single experiments by great scientists; our project is unusual in that we hope to uncover an entire trail of experiments. By analogy, just as Brunswik was critical of much psychological research for its reliance on single isolated stimuli, so also are we critical of much cognitive-historical explanation for its reliance on single experiments as the sole model system for analysis.

Ryan Tweney: tweney@bgnet.bgsu.edu

Seventeenth Annual Meeting of the Brunswik Society

Summerfield Suites by Wyndham, Lake Buena Vista, 8751 Suiteside Drive, Orlando, FL 32836
Organisers: Jim Holzworth (holz@uconn.edu) and Mandeep Dhami (mdhami@psyc.umd.edu).

Registration: Please register no later than November 1. You may register by email (info@brunswik.org), telephone (518-442-3850), or fax (518-442-3398). Registration fee (includes breaks and Friday lunch). Regular = \$50, Student = \$25 (student ID). Checks should be payable to "Brunswik Society" and sent to: The Brunswik Society, c/o Tom Stewart, 135 Western Ave. Milne 300, Albany, NY 12222

Accommodations: The meeting is held concurrently with the Psychonomic Society Annual Meeting and just before the Judgment and Decision Society meeting. These are in Coronado Springs Resort, Orlando, which is about 4 miles (less than 10 min taxi ride) from the location of the Brunswik Society meeting.

Summerfield Suites has reserved a few suites for participants who want to stay at the Brunswik meeting hotel. The Brunswik Society rate is \$109 for a one-bedroom suite and \$129 for a two-bedroom suite (two separate bedrooms and two baths). If you are interested, make a reservation by calling 407-238-0777 Monday to Friday, 9:00 to 5:00 pm, or by fax to 407-238-0778. email: mmillsaps@summerfieldsuites.com.

Why are some judges better than others?

Elise A. Weaver & Tom Stewart (Worcester, MA; Albany, NY, USA)

We have been studying individual differences in judgmental competence. We hypothesized that skill in multiple cue probability learning (MCPL) would predict accuracy in three judgment tasks that are familiar to most laypeople (baseball team success, apartment rents and used car prices).

We reasoned that the ability to learn to use multiple, fallible indicators is an important skill in making judgments, and that performance in an MCPL tasks would measure that skill. We also included tests of crystallized intelligence, fluid intelligence, short term memory (digit span) and coherence competence (as measured by problem solving tasks from the heuristics and biases literature). Our objective is to learn whether MCPL accounts for individual differences in judgmental accuracy that are not accounted for by intelligence or coherence competence. Results from a pilot study of 37 people were reported at the SJDM meeting last year. We have now collected data from a sample of 100 people and have begun the analysis.

Preliminary results suggest that fluid intelligence, short term memory, and coherence (probability rules) are predictors of judgmental accuracy. By the time we present our results at the Brunswik Society meeting, we will know whether MCPL ability is also related to accuracy.

Elise Axelrad Weaver: eaweaver@wpi.edu
Tom Stewart: t.stewart@albany.edu

AGENDA: Seventeenth Annual International Meeting of The Brunswik Society

Thursday, November 15, 2001

- 1:00 - 1:30 Late registration (Summerfield Suites by Wyndam)
1:30 - 1:40 Introductions
Welcome (Jim Holzworth & Mandeep K. Dhami)
1:40 – 2:40 Paper session 1: Applications I (Chair – Bob Wigton)
Leonard Adelman, Sheryl L. Miller & Cedric Yeo – Testing the effectiveness of perceptual interface components for supporting distributed team decision making
Louise Gunderson & Donald Brown - Using Derived Preferences to Predict the Targets of Computer Crime
Claudia Gonzales-Vallejo, Chad Muntz & Jessica Wilson – Not all judgments are equal: Exploring the differences between wholistic versus attribute-based judgments
2:40 - 3:00 Tea and coffee break
3:00 - 4:20 Paper session 2: Research design (Chair – Mike Doherty)
Yanlong Sun – Small samples, local anomalies and signal detection
Neil Bearden – Some results on learning in partially observable environments
Kris Canali & Jim Holzworth – Judgment analysis using efficient representative design
Mandeep K. Dhami – Evaluating the importance of representative design
4:20 - 5:10 Discussion session 1
Issue - On the matter of methods. What constitutes representative design? What does it mean in practice? (Chair – Jim Hogge. Discussants – Mike Doherty, Tom Stewart)
5:10 Adjourn
6:15 Evening dinner and entertainment at CityWalk, Orlando. Motown Café.

Friday, November 16, 2001

- 8:30 – 9:50 Paper session 3: Historical and theoretical analysis (Chair – Jeryl Mumpower)
Ryan Tweney & Elke Kurz – Making out a case for the experimenter's agency: Brunswik, Faraday, and the "inner laboratory"
Bernhard Wolf – Origins of the basic behavior-principles univocality and equivocality in Brunswik's system
Laura Martignon - Naive decision trees
Elise Weaver & Tom Stewart – Why are some judges better than others?
9:50 – 10:10 Tea and coffee break
10:10- 11:00 Discussion session 2
Issue - Cognitive Continuum Theory: Tests and Applications. What is the evidence for and against CCT? (Chair – Jim Holzworth. Discussants – Ken Hammond, Kathy Mosier)
11:00- 12:00 Paper session 4: Applications II (Chair – Tom Tape)
Kathy Mosier – Cognition in aviation in terms of intuition/analysis and coherence/correspondence
Gary McClelland & Barbara Fasalo - The importance of representative correlations in choice research
Mike Dekay – Conducting risk-ranking exercises as if the results really mattered
12:00 - 2:00 Buffet lunch at Summerfield Suites by Wyndam, and Jim Shanteau – Review of The Essential Brunswik
2:00 - 3:20 Paper session 5: Applications III (Chair – Clare Harries)
Paul Sorum & Tom Stewart – Examining the relation between physicians' diagnostic judgments and their treatment choices
Tom Tape, SB Konigsberg, MS Jackson, JR Bessmer, and DV O'Dell – How physicians decide whether to admit "low risk" chest pain patients
Liz Smith & Ken Gilhooly – Describing prescribing
Lisa Kath - Judgments about sexual harassment
3:20 - 3:40 Tea and coffee break
3:40 - 4:30 Discussion session 3
Issue - What is good judgment? (Chair – Paul Sorum. Discussants - Neal Dawson, Clare Harries)
4:30 - 4:45 New Investigator Prize (Awarded by Tom Stewart)
4:45 – 5:00 Farewell and meeting adjourned