



The Brunswik Society Newsletter

www.brunswik.org

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

November 2004, Volume 19

Contents:

2		Remembering Ledyard R Tucker by Tom Stewart
4	Abdollahi	Effects of Subliminally-Primed Bodily Positions and States on JDM
4	Adelman	Confirmation Bias in Complex Analyses
4	Balaji Rao et al	Application of Brunswikian Theory to Corrosion Damage Assessment of Reinforced Concrete Structural Members
5	Brehaut	How Decision Rules are Applied in Practice: A Clinical judgment analysis
6	Carvalho	The Use of JA and SJT in Legislative Research
6	Dhami et al	The Role of Representative Design in an Ecological Approach to Cognition
6	Dibadj	Further Analysis of the Lens Model Equation
7	Doherty	Sampling and Understanding the Ecology
7	Dunwoody	Nutrition Judgments, Intelligence Judgments, Uncertainty and Error
8	Garcia-Retamero et al	Compound Cue Processing within the Fast and Frugal Heuristic Approach
9	Good	Q-methodology and Probabilistic Functionalism: Exploring the Personal and Intellectual Links between William Stephenson and Egon Brunswik
10	Gunderson	Social Judgment Theory and the Stock Market
11	Hamm	Logistic Model Predictions for the Lens Model: How Many Steps?
12	Hammond	Judgment Under Uncertainty: Calls for Research
12	Harries	Advice-based Decision Making, Patient Choice & Doctors' Diagnosis
13	Harries	Training Occupational Therapists to prioritise appropriately
13	Harrison	Surgical Decision Making
13	Hastie and Kameda	Foraging for Majority and Plurality rules in Truth-Seeking Group Decisions
14	Hine et al	Cue Utilization and Harvest Decisions in Commons Dilemmas
14	Juslin	Brunswik's Methodological Fears are Music to Our Ears
15	Kirlik	Ecological Validity as a Mediator of Visual Attention Allocation
16	Koele and Ebels	Gender & Ethnic Background Influences on Personnel Selection Decisions
18	Lafon	MCPL among Children, Adolescents and Young adults: the Case of Inverse Relationships
19	Macbeth	Metacognition and Value Attribution
19	Mosier	Brunswik and Human Factors Research
20	Mumpower	Payoff Perception in Negotiations
20	Poses	Appropriate Judgment and appropriate medical treatment decisions
21	Shanteau and Gustafson	Solving the "User Override Problem": Why Do Experts Turn Off Decision Support Systems?
22	Sjodahl	Nursing attitudes and clinical inferences in psychiatry
23	Sorum	Do Physicians Recalibrate Patients' Pain for Functional Reasons?
24	Speekenbrink	Collaboration in NMCPL
25	Tschirhart	Quantity Judgment with Conflicting Information
25	Wearing	Research on Complex probabilistic worlds
26	Weaver	Travelling at an Interface
26	Wigton	Policy Recognition in Practitioners' Antibiotic Prescription
26	Yin and Rothrock	The Continuum between Compensatory and Noncompensatory Decision Strategies
28		20 th Annual Meeting of the Brunswik Society in Minneapolis

Ledyard R Tucker
1910 - 2004

Tom Stewart
University at Albany

Ledyard R Tucker (the “R” was his middle name, not an initial, so there should never be a period after it) was a brilliant psychometrician who studied with L.L. Thurstone and, after over a dozen years as the first director of statistical analysis at the Educational Testing Service, was Professor of Psychology and Educational Psychology at the University of Illinois from 1960 until his retirement in 1979. He made significant contributions in many areas of quantitative psychology and is known to Brunswikians for his refinement of the lens model equation.

Tucker (friends and some students called him “Tuck,” but I never felt comfortable enough for that) was extremely influential in shaping modern psychometric methods and quantitative psychology, but he is not well known outside of psychometrics because he tended to publish difficult papers on advanced topics. Few younger psychologists have heard of him, and he is virtually unknown outside of psychology. He deserves more recognition.

Although he published many important papers, his influence was even greater than his published work would indicate. He was at the center of a network of the world’s leading psychometricians. Every year, they came to Champaign-Urbana for several days to participate in his Factor Analysis Working Group, which included names like Gulliksen, Harmon, Horst, Jöreskog, Kaiser and Meredith. They met around a circular table, and a few students were privileged to observe.

I was one of the students in Tucker’s Measurement and Differential Psychology program at Illinois. In addition to a solid grounding in multivariate methods, he gave me an intense interest in and appreciation for the study of individual differences. I remember him as a kind, gentle, and somewhat frail man who smoked constantly. One semester, I was taking two seminars with him, and he was my dissertation advisor. As I watched him smoking and coughing in class, I hoped he would live until I finished my degree. That was 1969! Thirty-five years later he passed away at the age of 93.

Tucker’s students fondly remember his use of rubber balls and knitting needles to demonstrate the geometric interpretation of factor analysis. Although he had an astonishing record of never having an article rejected by *Psychometrika*, he was also legendary for his tendency to pull an unpublished paper out of his drawer on virtually any topic that would come up in discussions with him.

Tucker’s contribution to Brunswikian research was his 1964 paper on the lens model equation (LME). The editor of *Psychological Review* sent the Hursch, Hammond and Hursch manuscript to Tucker for review. Here is the rest of the story in Ken Hammond’s words:

The point of the story is that in his review Tucker came up with a simplified, and more coherent version of the LME, merely to help us along. I was so impressed with Tucker’s version of the LME, however, that I wrote to the editor and urged him to publish Tucker’s review. The editor agreed and asked Tucker for permission. Tucker was astonished (he

didn't find anything wrong with the manuscript) and rewrote his review to focus on the introduction of "G." And G has been there ever since, almost a half century. This was probably a unique event in the history of a relation between an author, a reviewer, and an editor.

Of course, Tucker's version of the LME is the one we use today, and his paper is frequently cited. I did not begin serious lens model work until I left Illinois and began working in Hammond's Center for Research on Judgment and Policy at Colorado. I never discussed the lens model equation with Tucker, although a couple of years ago I sent him my chapter on the LME from the *Essential Brunswik*, and got a very nice reply.

Another Tucker student, Lee G. Cooper of the UCLA Anderson School, adds:

He did the same thing in multidimensional scaling. A close reading of the Torgerson classic, *Theory and Methods of Scaling* (1958, p. 258) reveals the central equations for metric multidimensional scaling came from his derivation.

The most insistent demand of this gentle, understated man was that scientists must look carefully at their data. His genius was in providing us the mathematical and analytical tools to do so. Graphical rotation in factor analysis, multidimensional scaling, three-mode factor analysis, and powerful multivariate model for the understanding of the structure of individual differences were his lasting legacy. His legacy is also in the heartfelt appreciation of his students for depth and humanity of his mentoring.

Psychology, psychometrics, and the Brunswik Society have lost a brilliant mind and a wonderful teacher and colleague. There will be obituaries written by people who worked much closer with Tucker than I did. I've listed some web resources below for people who want to know more.

(I am grateful to Kenneth Hammond, Lee G. Cooper, and Richard Shikiar for their comments and contributions.)

References:

- Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Review*, 71(1), 42-60.
- Torgerson, Warren S. (1958), *Theory and Methods of Scaling*. New York: John Wiley and Sons.
- Tucker, L. R (1964). A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd. *Psychological Review*, 71(6), 528-530.

Links:

Neil J. Dorans, A Conversation with Ledyard R Tucker, ETS, Princeton, N.J.

<http://www.ets.org/research/dload/tucker.pdf>

A biography and interview that includes a photo of Tucker as I remember him

ETS press release on Tucker's death. <http://www.ets.org/news/04092401.html>

Exploratory Factor Analysis, A book manuscript by Ledyard Tucker and Robert MacCallum, <http://www.unc.edu/~rcm/book/factornew.htm>

Lubinski, D. (2004). Obituary: Lloyd G. Humphreys: Quintessential Scientist (1913–2003). *Intelligence*, 32, 221–226.

http://peabody.vanderbilt.edu/depts/psych_and_hd/smpy/HumphreysObit.pdf

This includes the story of Tucker's hiring at Illinois (p. 222)

Effects of Subliminally-Primed Bodily Positions and States on Judgment and Decision Making

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

Based on the findings of the embodiment and "modality-specific systems" theories, bodily positions and states can affect a wide variety of cognitive processes and behaviors (Barsalou, 1999; Barsalou, Niedenthal, Barbey, & Ruppert, 2003). An interesting relevant question for both embodiment and judgment and decision making researchers would be whether the embodiment effects could be induced nonconsciously through subliminal priming and in this way produce a judgment and/or decision which might be considered "nonconsciously-made."

To investigate this hypothesis, two experiments were conducted. In Exp 1, 89 participants were (were not) subliminally primed with various pictures each illustrating different bodily positions and states involving different emotions. To rule out the possible direct effect of emotions, the stimuli were designed not to suggest any emotion-specific facial expressions. For example, one picture showed a person standing with fully open and extended arms and having an emotionally-neutral face. Then, participants were required to make a series of consumer judgments and decisions (e.g., evaluating and buying a DVD player). In Exp 2, 96 participants were (were not) subliminally primed with words representing the bodily positions and states involving different emotions. Here too, participants made similar judgments and decisions. Results revealed that in both experiments, the primed bodily positions and states affected participants' judgments and decisions. That is, they tended to nonconsciously match their judgments and decisions to the emotions produced by the "observed" bodily positions and states. These findings may indicate that there is a direct relationship between "implicit embodiment" and judgments and decisions.

References

- Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-609.
- Barsalou, L.W., Niedenthal, P.M., Barbey, A., & Ruppert, J. (2003). Social embodiment. In B. Ross (Ed.), *The Psychology of Learning and Motivation*, Vol. 43 (pp. 43-92). San Diego: Academic Press.

Confirmation Bias in Complex Analyses

Len Adelman
George Mason University
ladelman@gmu.edu

Most research investigating the confirmation bias involves abstract experimental tasks where subjects draw inferences from a few items of evidence. These tasks are not representative of complex analysis tasks characteristic of law enforcement investigations, intelligence and financial analysis, where there is substantial evidence and the evidence items vary in interpretability, reliability and validity. Colleagues (Paul Lehner, Brant Cheikes, and Mark Brown) at the MITRE Corporation and I performed an experiment investigating the confirmation bias for a representative, complex analysis task. Results indicated that participants tended to overweight the information that was consistent with the hypothesis they considered most likely. Information misinterpretations were rare; for example, saying that information supported a hypothesis when it didn't. Rather, participants overweighted information consistent with prior believed hypotheses and underweighted the information that wasn't.

Application of Brunswikian Theory to Corrosion Damage Assessment of Reinforced Concrete Structural Members

K. Balaji Rao
Scientist¹
balaji@sercm.csir.res.in

M.B. Anoop
Scientist¹

N. Lakshmanan
Director¹

S. Gopalakrishnan
Advisor [Management] and Director Grade Scientist¹

T.V.S.R. Appa Rao
Emeritus Scientist [CSIR] and Formerly Director¹
Structural Engineering Research Centre, Chennai, INDIA

The corrosion of reinforcement in concrete is an issue of major concern, as billions of dollars are spent every year for the maintenance/repair of reinforced concrete structures damaged due to corrosion. The decision regarding maintenance/repair of the damaged structures should be based on a rational assessment of the damage. The authors have made an attempt to

apply Brunswikian theory (Wolf, 2000) for taking into consideration the uncertainties associated with human mental process in making judgment regarding corrosion damage state. A multi-level lens model for corrosion damage assessment of reinforced concrete structural members is proposed (Balaji Rao et al, 2004). The cognitive process in judgment is modelled using the probabilistic mental model (PMM) proposed by Gigerenzer et al (Gigerenzer, Hoffrage & Kleinbolting, 1991). The achievement of each expert making judgments is determined using the generalised linear model, and the over- or under-confidence limits associated with the experts are determined based on the judgments on a number of baseline cases. In the present study, it is assumed that the linear regression analysis has been carried out and the values of achievement index have been obtained for all the experts making judgments. It is also assumed that the experts have similar tendencies in viewing the environment and making judgments (viz. all having positive correlations). The judgments made by different experts are combined together by taking a weighted average (the weights are determined based on the achievement index values of the experts, reflecting the correlation between the judgment and the environment). The over- or under-confidence associated with the corrosion damage state is also determined, which is useful as a measure of the confidence that can be put on the corrosion damage state obtained based on experts' judgments. One of the highlights of the methodology is that those involved in making judgments are considered to be rational, rather than classifying them as experts and non-experts, and the performance of each expert is determined by computing the achievement index. This is in line with the current thinking in risk perception and risk communication (Reid, 1999), and would help in creating a more effective knowledge-based system for damage assessment. The studies are being continued also including the evaluation of experts involved in making judgments.

References

- Balaji Rao, K., Anoop, M.B., Lakshmanan L., Gopalakrishnan, S., & Appa Rao, T.V.S.R. (2004). Risk-based remaining life assessment of corrosion affected reinforced concrete structural members. *Journal of Structural Engineering*, 31(1), 51-64.
- Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98(4), 506-528.
- Reid, S.G. (1999). Perception and communication of risk and the importance of dependability. *Structural Safety*, 21(4), 373-384.
- Wolf, B. (2000). Processes of constructing judgments and actions by competent individuals with respect to object orientation: Programmatic ideas in the tradition of Brunswikian thoughts. The Brunswik Society, Essay #7, <http://www.brunswik.org>

How Decision Rules are Applied in Practice: A Clinical Judgment Analysis

Jamie Brehaut
Ottawa Health Research Institute,
Ottawa, Ontario, Canada
 jbrehaut@ohri.ca

Clinical decision rules (CDRs) are algorithmic decision tools designed to help health care practitioners make specific diagnostic or treatment decisions (e.g. do I need to order ankle radiography for this patient?). The benefits of widespread use of appropriately developed CDRs have been discussed in the literature, and include reduced health care costs and practice variation, simpler, more evidence-based physician decisions, and reduced wait times and unnecessary procedures for patients. While some CDRs are widely disseminated and reported to be used (e.g. the Ottawa Ankle Rules are reportedly used by nearly 100% of Canadian Emergency physicians), there exist no data beyond self-report to address the issue of how these rules have actually been incorporated into physicians' decision-making processes. We (with Bob Wigton, Ian Stiell and others) are conducting a clinical judgment analysis of Canadian emergency physicians' use of the Ottawa Ankle Rules via postal survey, to examine whether physician judgment policies will match that of the rule they report using. A previous survey showed that many of these physicians report using the Ankle Rules non-exclusively; that is, even in situations where it would be appropriate to govern their decision by the rule alone, other non-rule factors are considered. The judgment analysis will allow us to determine what kinds of non-rule factors (i.e. cues that indicate the rule should not be used, cues related to fracture but that add no additional predictive value over the rule, cues that are unrelated to fracture) enter into the decision. Two hypotheses to be tested include 1) The extent to which decisions are rule-based will be consistent among identifiable subgroups of physicians (e.g. those who say they use the rule alone vs. those who say they

consider other factors); 2) Use of rule-based cues in decision-making will be more consistent across physicians than will use of non rule-based cues. We have just completed data collection and are in the process of data entry and analysis. I hope to present these findings at next year's Brunswik meeting.

The Use of Judgment Analysis and Social Judgment Theory in Legislative Research*

Marcio A. Carvalho #
University at Albany, USA
macarvalho@fgv.br

For many years, researchers have been trying to understand the decision-making process within legislative bodies. Their main objective is to try to forecast future roll-calls, to better influence legislative members and to improve the legislature's internal process. Since most of these studies do not differentiate between judgment and decision-making, they use the final outcome of a legislator decision – his or her vote – as the main tool for their analysis, and the main indicator of future behavior. The objective of my research is to propose and test a new methodology in legislative studies, based on Social Judgment Theory (SJT) that could allow researchers to understand the judgment process behind legislative decisions. In order to do that, a demonstration study was performed to simulate the use of judgment analysis in an imaginary House of Representatives. The results show that the use of SJT tools in legislative research is feasible. A broader study is being conducted in Brazil aiming to use the proposed methodology with real representatives. This study uses three methods that aim to create a more consistent judgment analysis study: interviews, analysis of existing data, and regression analysis of judgments. The simulations study and the initial results from the study being done in Brazil show that the use of this method could add to the fields of decision making and legislative research, creating more robust models that could better represent the behavior of members of legislative bodies.

* This research has being conducted thanks to a grant from FGV/EBAPE in Brazil, and the guidance of Professor Thomas Stewart (Center for Policy Research – SUNY/Albany)

Marcio A. Carvalho is an assistant professor at Fudacao Getulio Vargas (FGV/EBAPE) in Brazil, PhD candidate in Public Administration at the Rockefeller

College of Public Administration and Policy, State University of New York at Albany (SUNY/Albany), masters in Public Administration (SUNY/Albany) and Political Science (UnB/Brazil).

The Role of Representative Design in an Ecological Approach to Cognition

Mandeep Dhami
University of Victoria, BC, Canada
mkdhami@uvic.ac.uk

Ralph Hertwig
University of Basel, Switzerland

Ulrich Hoffrage
Max Planck Institute for Human Development, Germany

Half a century ago, Egon Brunswik argued that psychological processes are adapted to the uncertain environments in which they evolved and function. He proposed the method of representative design to capture these processes and advocated that psychology be a science of "organism-environment" relations. Representative design involves randomly sampling real stimuli from the environment or creating stimuli in which environmental properties are preserved. Thus, it departs from the tradition of systematic design endorsed in research texts. In this paper (the present abstract is taken from a manuscript published in *Psychological Bulletin*, Vol. 130, pp. 959-988), we review the development of representative design, examine its use in the field of judgment and decision-making, and demonstrate the effect of design on research findings. We suggest that some of the practical difficulties associated with representative design may be overcome with the use of modern technologies. Finally, we discuss the importance of representative design and sampling in several areas of psychology, and the implications of this method for ecological approaches to cognition.

Further Analysis of the Lens Model Equation

Hassan Dibadj
University at Albany, USA
hd7554@albany.edu

My analysis of the interrelationships among the different parameters of the lens model equation is almost complete. Statistical properties of these parameters have been also studied. I am also completing analysis of the coefficient of the reliability of

information acquisition in the expanded lens model. This analysis will help any Brunswikian researcher to know how to interpret and use the lens model equations.

Sampling and understanding the Ecology

Mike Doherty
Bowling Green State University, USA
mdoher2@bgsu.edu

Yaakov Kareev and his colleagues recently published a number of provocative papers asserting that the environment, mathematically conceived, favors small samples over large ones for the intuitive assessment of correlation. They have also argued for the adaptiveness of such a small sample advantage. The argument for this "small window" effect, though not framed in a Brunswikian fashion, is deeply Brunswikian, in that it is predicated upon a careful analysis of the structure of the ecology, which, as we all know, Brunswik believed is propaedeutic¹ to the investigation of behavior in that ecology. Richard Anderson and I have been investigating this "small window" effect with extensive simulations and also behavioral investigations. A forthcoming *Psychological Review* will have two simulation-based analyses of the small window hypothesis, one by Peter Juslin and one by Richard Anderson, myself and two grad students (Neil Berg and Jeff Friedrich, as well as a rejoinder by Yaakov Kareev.

We also have preliminary behavioral data (60 Ss) on the inference of correlation from arrays of bar graphs, analyzed via signal detection theoretic techniques. The data show a small nonsignificant difference in hit rate favoring samples of 3 over samples of 15, but significant ($p < .01$) differences in false alarm rate and hit rate minus false alarm rate favoring samples of 15 over samples of 3.

A chapter (Doherty, M. E., & Tweney, R. D. entitled Reasoning and task environments: The Brunswikian approach, has just been published in the book, *Psychology of Reasoning: Theoretical and Historical Perspectives*, Psychology Press, edited by Ken Manktelow & M. C. Chung. I gave a talk with the same title at the 5th International Conference on Thinking in July, in Leuven, Belgium. I haven't seen the book yet, but I know that it's out.

1. Providing introductory instruction (Dictionary.com) – Ed

Nutrition Judgments, Intelligence Judgments, Uncertainty and Error

Philip T. Dunwoody
Assistant Professor of Psychology
Juniata College, Good Hall, USA
<http://faculty.juniata.edu/dunwoody/>
drdunwoody@yahoo.com

I am settling into a new position at Juniata College, a small liberal arts college in the mountains of central Pennsylvania. In addition to the ongoing projects I described in last year's newsletter, I have two main lines of research that I am developing.

The first involves the development of a simulation to study how people use information from the Nutrition Facts Panel (the federally mandated label on the back of all food products). The simulation is being designed and developed by myself and two colleagues, Esmail Bonakdarian (currently at Oberlin College) and Christopher Marino (currently at Tri-County Community College), and we plan to start testing it and collecting data in Spring '05. Once the simulation is operational, the first phase of the research will be to recruit SMEs to judge all products based on their overall nutritional value and attempt to model their judgment profile. Phase two: Once we have a satisfactory view of how experts perform this task we will then have college students make similar judgments and assess their competency using the SMEs as the criterion. Phase three: We will then try and make improvements to the Nutrition Facts Panel with the goal of getting novice users (college students) to make judgments more similar to SMEs. We hope to be able to improve judgment performance merely by changing the information presentation (no additional nutrition education will be provided). This will be a multi-year project and I expect to report more specifics about phase one in next year's newsletter.

The second line of research that I am developing is non-experimental (at least for now) and largely inspired by recent events. I am starting to research social policy judgments with the goal of contributing to our understanding of how our intelligence community (IC) and policy makers assess threat. Threat assessment of a group (terrorist organization) or state (rogue-state or otherwise) involves the integration of multiple fallible indicators (such as intelligence). Recent years have seen a number of judgment errors both within the IC and our public leadership. Ken Hammond and I have co-authored a paper (currently under review) applying the Taylor-Russell diagram to the policy of pre-emption which frames the difficulties of a predominantly pre-

emptive strategy in terms of multiple fallible indicators, irreducible uncertainty, and the duality of error. I will be discussing this paper at this year's Brunswik meeting.

I believe that Brunswikian theory can help contribute to our understanding of judgment error within the IC and aid policy makers as consumers of intelligence in a number of key ways. Namely, we should attempt to estimate the empirical accuracy of intelligence analyses when no clear criterion is available. Such estimates could be used by policy makers as estimates of the ecological validity of a particular intelligence estimate. I will discuss some specific ways that this may be accomplished at this year's Brunswik Society meeting in the symposium entitled, "Could the Brunswik Society Fix the Broken Parts of the U.S. Intelligence System?"

Compound Cue Processing within the Fast and Frugal Heuristic Approach

R. García-Retamero^{1,2}
U. Hoffrage²
A. Dieckmann²
M. Ramos³

¹University of Granada, Spain

²Max Planck Institute for Human Development, Berlin, Germany

³University of Jaén, Spain

rretamer@mpib-berlin.mpg.de

Most of the decisions we make in everyday life have to be fast and based on little information. How else are we to decide which of several articles would be worth our time to read next? How do we make such decisions? One recent approach, promoted by the ABC Research Group, suggests that people use *fast and frugal heuristics* in these situations, that is, simple, but nevertheless often accurate decision rules (Gigerenzer, Todd, & the ABC Research Group, 1999). One example for such a fast and frugal heuristic is *Take The Best* (TTB).

Simple heuristics like TTB do not take interactions between cues into account to predict a criterion. However, evidence in several research areas suggests that people can and do process combinations of cues (i.e., compound cues) using configural strategies (Edgell, 1993). Along these lines, we propose the *Take The Best Configural Cue* heuristic (henceforth, TTB Configural; García-Retamero, Dieckmann, Hoffrage, & Ramos, 2004) as a one-reason decision-making

heuristic that takes compound cues into account. To be fast and frugal, this heuristic relies on people's causal assumptions (i.e., the perception that the component cues of the compound act through a common causal mechanism) as a precondition for processing compound cues as configurations that are subsequently included in the hierarchy of cues ordered according to validity. In short, intuitive causal assumptions allow the decision maker to deal adaptively with the infinite number of combinations of cues that appear in a particular environment by directing them to the compound cues that are relevant.

We ran two series of experiments to find out whether the TTB Configural heuristic was an adequate model for predicting participants' decision behavior (García-Retamero, Dieckmann, Hoffrage, & Ramos, 2004; García-Retamero, Hoffrage, Dieckmann, & Ramos, 2004). Participants were provided with an environment that included a valid compound cue, which was constructed using the logical XOR rule (the eXclusive OR) and its opposite, the nonXOR rule (Experiments 1 and 2 in the first series of experiments), or using the AND and the OR rule (Experiments 1 and 2 in the second series of experiments). In a control condition, participants were given a simple environment where there was no valid compound cue (Experiment 3, both in the first and second series of experiments). In each experiment, we also manipulated participants' causal assumptions about the cues in the environment (i.e., the causal mental model with which they approached the task).

The results indicated that a high percentage of participants used the TTB Configural heuristic if there was a valid compound cue in the environment (regardless of which logical rule constituted this compound), and if the instructions suggested that its components act through a common causal mechanism, thereby prompting a configural causal model of the environment. However, when cues were said to act through different causal mechanisms, or when no information about the causal mechanism was provided, the TTB heuristic was used by a high percentage of participants even when there was a highly valid compound cue in the environment. A high percentage of participants also used the TTB heuristic when no valid compound cue existed. These results showed that, through the proposed heuristic, fast and frugal heuristics can exploit an even wider range of information structures to make adaptive decisions.

References

- Edgell, S. E. (1993). Using configural and dimensional information. In: N. J. Castellan (Ed.), *Individual and group decision making* (pp. 43-64). Hillsdale, NJ: Erlbaum.
- García-Retamero, R., Dieckmann, A., Hoffrage, U., & Ramos, M. (2004). *Compound stimuli processing within the fast and frugal heuristic approach in nonlinearly separable environments*. Manuscript submitted for publication.
- García-Retamero, R., Hoffrage, U., Dieckmann, A., & Ramos, M. (2004) *How simple heuristics can process compound cues in linearly separable environments*. Manuscript submitted for publication.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Q-methodology and Probabilistic Functionalism: Exploring the Personal and Intellectual Links between William Stephenson and Egon Brunswik

James Good
Centre for the History of the Human Sciences
And Department of Psychology
University of Durham
j.m.m.good@durham.ac.uk

This paper is part on an on-going project on the life and work of William Stephenson. It employs material from the William Stephenson Archive housed in the Western Historical Manuscripts Section of the Ellis Library, University of Missouri-Columbia and personal correspondence and papers that the Stephenson family have allowed me access to. The UK Leverhulme Trust funded some of the research on which the paper is based.

In the paper I explore some personal and intellectual links between William Stephenson and Egon Brunswik. I note some striking commonalities in their lives and ideas and some interesting parallels in the reasons for the lack of acceptance of their ideas. The paper begins by outlining some aspects of their personal relationship, drawing on some of Stephenson's personal correspondence. The two men's lives were parallel in a number of respects. They were born less than a year apart. Brunswik studied engineering (Stephenson physics) and finally took his Ph.D in 1927

(Stephenson 1929). Both made extensive use of correlation in their measurements.

Stephenson and Brunswik first met in London in 1929 – most probably at University College London where Stephenson was completing doctoral studies in psychology with Charles Spearman. Stephenson emigrated to the United States in 1948 and renewed his acquaintance with Brunswik when he paid several visits to Berkeley in the late 1940s. He was also a Visiting Professor at Berkeley in 1951-2.

Stephenson incorporated the concept of representative design and the notion of ecological universes in his second book *The Study of Behavior: Q-Technique and its Methodology* (Stephenson, 1953). Stephenson suggested that the former concept applied to attitudes, opinions, and beliefs: [a] “sample of these statements will be called ‘representative’ if it in some way samples the behavior adequately, that is, without favoring some parts to the exclusion of others and without omitting whole sections of the universe of possibilities” (Stephenson, 1953: 222).

Brunswik makes reference to Stephenson's work in his *Systematic and Representative Design of Psychological Experiments* (1947) but clearly wishes to distinguish his own use of correlation as dealing with functional problems from Stephenson's “inverted correlation” technique. Here, as Steve Brown has pointed out (Brown, 1995), Brunswik is a victim of the widely held view that R and Q represent reciprocal features of one and the same data matrix. Stephenson's use of correlation thus was represented in Brunswik's comprehensive Table 2 as a branch of differential psychology to distinguish it from his own functional and ecological psychology. The 1956 revision of the 1947 monograph does suggest that Brunswik was becoming better acquainted with Stephenson's work as it employs Stephenson's notion of intra-individual differences (Brunswik, 1956).

While examining the Stephenson archives I discovered that Brunswik had been the University of Chicago Press reader for Stephenson's 1953 book. The archives included Brunswik's report and a personal letter sent at the same time by him to Stephenson. These letters are now on the Brunswik web pages as is a very helpful contextualising note by Kenneth Hammond (Hammond, 2002).

The paper continues by noting that although Brunswik thought highly of Stephenson's achievements (“I have just finished reading your stupendous manuscript...I am adding here some comments for you personally” (Brunswik, May 26, 1952), he had reservations about

some aspects of Stephenson's position – especially in relation to his use of the idiographic-nomothetic distinction. I suggest that while Brunswik may have misunderstood some aspects of the significance of Stephenson's work (with regard to the distinctiveness of Q-methodology), it also the case that there are features of Brunswik's work that Stephenson failed to comprehend fully – especially his stance on the idiographic/nomothetic distinction.

The paper concludes by referring to David Leary's (1987) discussion of three inter-related factors which he believed had contributed to the misunderstanding and incomprehension from which Brunswik's psychology suffered. These factors were: first, the constitution of Brunswik's psychology - involving a unique blend of traditions, intellectual and scientific; second, as a consequence of his move to the United States, the loss of an audience that shared much of his intellectual background and a professional context in which his work on perception was seen as research worthy and important; third, Brunswik's style of communication, frequently involving philosophical and historical analyses, and enthusiastic discussions of new developments in other sciences and disciplines which were not always matched by his students' (and colleagues') ability to see the significance of what he was talking about. I suggest that these factors can also be seen as applying to the unfavourable reception of much of Stephenson's work. I surmise that had their work been received more favourably, the current landscape of the psychological sciences might have been very different indeed.

References

- Brown, S. (1995). Egon Brunswik and William Stephenson's Q Methodology. *Brunswik Archives*, Tuesday 10 January 1995.
- Brunswik, E. (1947). *Systematic and Representative Design of Psychological Experiments*. Berkeley and Los Angeles: University of California Press.
- Brunswik, E. (1952). Unpublished letter from Egon Brunswik to William Stephenson. May 26, 1952. Western Historical Manuscripts Section, Ellis Library, University of Missouri-Columbia.
- Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. Berkeley: University of California Press.
- Hammond, K.R. (2002). Brunswik letters. *Brunswik Archives*, Tuesday, 10 September 2002.
- Leary, D.E. (1987). From act psychology to probabilistic functionalism: The place of Egon Brunswik in the history of psychology. In M.G. Ash & W.R. Woodward (Eds.) (1987), *Psychology in Twentieth-Century Thought and Society*. Cambridge, England: Cambridge University Press (pp. 115-142).
- Stephenson, W. (1953). *The Study of Behavior. Q Technique and its Methodology*. Chicago: University of Chicago Press.
- Paper presented at the 20th Annual Conference of the International Society for the Scientific Study of Subjectivity, University of Georgia, Athens, Georgia, 23-25 September 2004.

Social Judgment Theory and the Stock Market

Louise Gunderson
Gamma-two, Inc.
lgunders@gamma-two.com

About this time last year, a client asked us to investigate a hedging strategy called 'pairs trading'. We used social judgment theory to develop a predictive model of the stock market. Pairs trading uses the following steps. Find two stocks that have a high correlation coefficient over the long term and a low correlation coefficient over the short term. Open a position by buying the undervalued stock and selling short the overvalued stock. When/if the short term correlation rises close out the position. This strategy requires that highly correlated stocks, when they have become uncorrelated, regain a high degree of correlation, at least most of the time.

We structured this as a social judgment theory question. Let us assume that the price of a stock is the result of a number of individual judgments about that stock. As this is a wildly stochastic and dynamic environment, any feature will have more or less functional validity, but individuals hold tightly to their specific set of features. (This is demonstrated by the interesting artifact of well respected stock pickers using astrology to make their selections.) We need to further assume that some features may act as 'shock terms' and lose value in the judgment process over time. If these two assumptions are true, then we would expect to see the following case. Two stocks are perceived to have approximately the same relative value (and therefore a high correlation). A feature of one of the stocks is affected by a rumor (or other event, such as a

news article). This affects the perceived value of that stock, thus causing the two stocks to have a low short term correlation. The value of the rumor diminishes over time, causing the perceived value of the stock to revert to its original value, thus causing the two stocks to be highly correlated again.

The next question is “how often does this happen?” We then proceeded to test the hypothesis by constructing a test harness in which the correlations for all the stocks in the three major exchanges (AMEX, NASDAQ, and NYSE) were calculated for the period 1/1/2004 – 5/31/2004. All of the stock pairs that had a 500 day correlation of greater than 0.90 and a 20 day correlation of less than .5 were identified. This happened 32,978 times in 16,506 pairs. The number of days until their short term correlation became greater than 0.9 was calculated. Only 3% of the stocks failed to regain a high short term correlation in 60 days.

The resulting software has been licensed to the client and we are waiting to see how much more research it will fund.

Revised November 13, 2004.

Logistic Model Predictions for the Lens Model: How Many Steps

Robert M. Hamm, PhD
*Clinical Decision Making Program
Department of Family and Preventive Medicine
University of Oklahoma Health Sciences Center, USA
robert-hamm@ouhsc.edu*

I have been working with Tom Stewart and Tom Tape on the Lens Model using logistic rather than linear regression (Cooksey, 1996). We were concerned about some anomalies in the data – the lens model did not “add up” when applied to categorical data. Before Tom Stewart figured out our problem (which we will present at the Brunswik Society meeting in Minneapolis), I got interested in the problem of fitting categorical data: categorical predictions are so far off when they are wrong, surely there is some information lost.

Logistic regression typically is useful for predicting events: yes or no, 1 or 0. Thus our environment data Y_e might be whether an event happened or did not happen, and our judgment or subjective data Y_s might be the judge’s prediction whether the event happened. The logistic regression produces two “predictions” – a continuous variable representing the probability the event happened, and a categorical best guess of

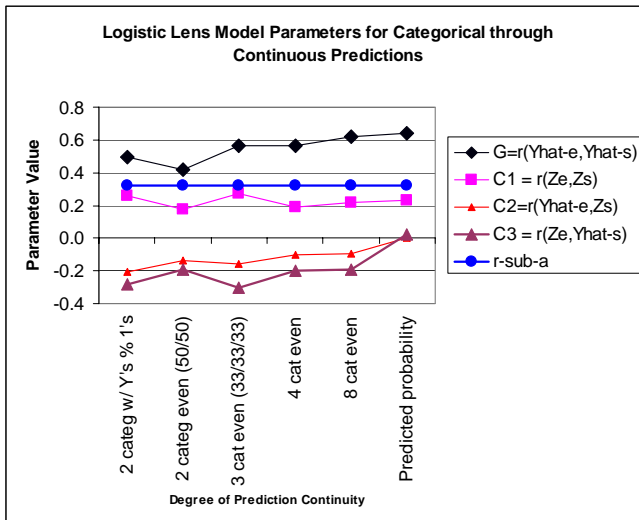
whether it happened or not. This guess is produced by setting a threshold on the continuous variable, and any value higher than the threshold is a prediction of a 1, else of a 0.

Which Lens Model would be best: one based on the continuous prediction (for both the environment model and the judge model), or one based on the yes/no prediction? Using the continuous logistic regression prediction ought to preserve the most information; but on the other hand if the Y_e and Y_s variables themselves are categorical shouldn’t we model them as such? I explored this using a convenient data set. (It had 240 judgments, each by a different person: Y_e = whether he got screened for prostate cancer; Y_s = his intention to get screened; the input X’s were various beliefs about prostate cancer screening.)

I constructed categorical prediction variables with different numbers of “steps” – 2, 3, 4, and 8. These were done by rank ordering the continuous logistic model prediction (a probability), and picking cutoffs that produced equal halves, thirds, and so on. I also considered a dichotomous prediction with cutoff equal to the proportion of 1’s in the Y_e or Y_s (which was not 50%), and the continuous probability variable. In this way, 6 logistic regression Lens Models were produced from the same data set, each using the same structure for the environment and judgment models. The Figure shows the correlation parameters but not the adjustment parameters (which are ratios of products of standard deviations). Y_e is the environmental criterion, \hat{Y}_e the model’s prediction, and Z_e its residual.

As the predictions got more steps, the knowledge parameter, G , increased, while the unmodeled knowledge parameter, $C1$, stayed about the same. The achievements $r_a = r(Y_e, Y_s)$ are the same, because Y_e and Y_s do not change. The crossed terms $C2$ and $C3$, which are negative correlations, increased toward 0. (These have to do with the relation between the predicted part of one side of the lens, and the unpredicted part of the other side.) Interestingly, while G had its biggest increase between 2 and 3 categories, the biggest decrease in the size of $C2$ and $C3$ occurred between 8 categories and the continuous prediction.

I expected the continuous prediction logistic models to perform better in all ways. Indeed, G increased as the number of categories in the prediction increased. However, even 3 categories produced a fairly accurate G (think of it as “yes, not sure, no”, for example). Of the 2-category prediction models, G was higher when the % of Y-hats that were 1 equaled the % of Y’s that were 1, rather than 50%.



The extra parameters C2 and C3 of the continuous prediction model were negligible, which is good because they are so difficult to explain. As I reported at last year's Brunswik Society meeting, the relative weights of the continuous prediction version of the logistic Lens Model are very similar to those of the linear Lens Model. It is possible, however, that the weights of a logistic Lens Model with a 2 category prediction might be quite different from those of the linear Lens Model.

Judgment Under Uncertainty: Calls for Research

Ken Hammond
 Boulder, Colorado, USA
 Kenneth.Hammond@colorado.edu

Last year I reported that I was working on a book ms; progress, I said, was "somewhat uneven". I am happy to report that progress is now "even". That is, yesterday I sent the publisher a ms that I am happy with, and that I am pretty sure my editor will be happy with. I have given up trying to write a book that will appeal to the airport reader and have settled on a ms that will be accessible to the well-educated reader as well as to the professional. I am reasonably confident that I have achieved this goal, and that you won't get a similar letter about "giving up" next year.

I should add one item of interest. The breakdown of the military intelligence agencies, and the (almost certainly) futile effort by the government to repair matters by administrative reorganization, motivated me to add a chapter pointing out that the real difficulty lies in our less than perfect knowledge about judgment under uncertainty. My chapter contains a suggestion

for a remedy: a "New Manhattan Project" for a strong new unprecedented research effort on the part of the granting agencies (NSF?) to further our understanding of this critical feature of sustaining the future of civilized society. This not the place to expand on this suggestion, but I will be happy to discuss it with interested colleagues.

Advice-based Decision Making, Patient Choice and Doctors' Diagnosis

Clare Harries
 University College London
 clare.harries@ucl.ac.uk

As usual, I am pursuing a few disparate strands of research. Here is a selection. Brunswikian principles are more bubbling under my interpretation and understanding of topics, than ruling the methodology. Several students are following up work on advice-based decision making – none want to study it in its natural ecology, but one is at least considering the statistical characteristics of naturally occurring advice. A project on risk communication and trust, with Nigel Harvey and Matt Twyman (University College London), based on risk of death figures as we can find them, explores the factors affecting advice giving and taking, and self-other judgment differences.

Speaking of self-other judgment differences, I'm involved in research-based discussions on facilitating patient involvement in medical decision making. Self-other judgment differences are of course key in an environment in which historically decisions have been made by an 'other' (health professional). My main role in these discussions has been to question the automatic assumption of the goodness of patient involvement in decision making. It may, or may not be the best way to achieve their goals. However, where a patient does not express a preference, a health-care professional is left to use the information they have about the patient as a best guess as to what their wishes might be. This work relates to the work on physicians' vicarious substitution of cues that I've written about in previous newsletters and that I'm still redrafting! Finally, as if I was shying away from an achievement perspective altogether, I and colleagues at the University of Birmingham (Olga Kostopoulou and Brendan Delaney) have won a grant to examine Diagnostic 'Errors' in General Practice. But the labelling is that of the grant awarding body, not ours. One of our central points of interest is the role that exposure to different types of cue environment, during six monthly

hospital placements, has on doctors' subsequent interpretation of patterns of cues in General Practice.

Training Occupational Therapists to Prioritise Appropriately

Dr. Priscilla Harries
MSc Course Leader in Occupational Therapy,
London, UK
priscilla.harries@brunel.ac.uk

I have completed my PhD this year using judgment analysis as the methodology. My thesis is entitled Occupational therapists' judgment of referral priorities: expertise and training. I was formally supervised by Professor Ken Gilhooly and given excellent advice and support by both Dr. Clare Harries and Prof Jim Shanteau.

The studies in my thesis describe how experienced occupational therapists' referral prioritisation policies were used to successfully educate novices. For occupational therapy services, where service demand far exceeds service availability, skill in referral prioritisation is essential. 40 British occupational therapists' referral prioritisation policies were modelled using judgment analysis. Individuals' prioritisation decisions were regressed onto 90 referral scenarios to statistically model how referral information had been used. It was found that the reason for referral, history of violence and diagnosis were most important. The occupational therapists' capacity for self-insight into their policies was also examined by comparing statistically modelled policies derived from their behaviour with their subjective view of their cue use. Self-insight was found to be moderate (mean $r = 0.61$).

A Ward's cluster analysis was used on the statistically modelled policies to identify if subgroups of therapists had differing referral prioritisation policies. Four clusters were found. They differed according to several factors including the percentage of role dedicated to specialist occupational therapy rather than generic work. The policies that led to more of an occupational therapy role were found to give particular importance to the reason for referral and the client's diagnosis. The occupational therapy professional body supports this latter method of working as it has recommended that occupational therapists should use their specialist skills to ensure clients' needs are met effectively. Therefore the policies that focussed on clients' occupational functioning were used to train the novices. Thirty-seven students were asked to prioritise a set of referrals before and after being shown graphical and descriptive

representations of the policies. Students gained statistically significant improvements in prioritisation. Students' pre-training policies were found to be those of generic therapists; a method of working that has been found to be leading to reduced work satisfaction and burnout (Craig et al.1998b). The training is therefore needed to ensure undergraduate occupational therapy students develop effective referral prioritisation skills. This will help to ensure that clients' needs are met most effectively and work stress is reduced.

The research has now taken another step forward as I have been funded to take the expert policies and use them to develop an on-line World Wide Web training package. This will be used to train novices in the task of referral prioritisation.

Surgical Decision Making

Michael Harrison
At present on sabbatical in the UK, in
Aberdeen, Scotland.
michael.harrison@auckland.ac.nz

Myself and a 'summer student', Sea-Young Yoon, examined the decision-making that surrounds the problem of whether to transfuse blood to a patient during surgery. We had to determine how pathologies ranked in the decision process and how contextual qualities affected the anaesthetist, such as peer pressure, surgeons' wishes and the likelihood of ongoing haemorrhage. It was possible to create a model that could improve the consistency of decision making. Clinical practice varies so widely that reducing this variation may have some benefits.

Foraging for Majority and Plurality rules in Truth-Seeking Group Decisions

Reid Hastie
University of Chicago, USA
reid.hastie@gsb.uchicago.edu

Tatsuya Kameda
Hokkaido University, Japan

Reid Hastie (University of Chicago) and Tatsuya Kameda (Hokkaido University) have been using the Brunswikian framework to evaluate the adaptive success of various group decision rules in a simulated

foraging task. In the simulation and in some behavioral experiments "subjects" seek resources in locations in an uncertain environment. Their judgment policies are simulated or "captured" using traditional lens model, multiple cue probability learning frameworks and statistics. A draft paper titled, "The Robust Beauty of Majority Rules in Group Decisions" is available from the authors and forthcoming in *Psychological Review*. The abstract runs as follows: How should groups make decisions? There is a long history of evaluations of social choice rules based on analytic tests of logical coherence, decisiveness, and conformity to the ideals of democratic social welfare. We shift the basis of evaluation from preferential conflict resolution to adaptive accuracy in choosing the mutually most beneficial alternative. We provide an original evaluation of nine group decision rules based on their adaptive success in a simulated test bed environment. When the adaptive success standard is applied, the majority and plurality rules fare quite well, performing at levels comparable to much more resource-demanding rules such as an individual judgment averaging rule. The plurality rule matches the computationally demanding Condorcet Majority Winner that is standard in evaluations of preferential choice. We also test the results from our theoretical analysis in a behavioral study of nominal human group decisions and the essential findings are confirmed empirically. The conclusions of the present analysis support the popularity of majority and plurality rules in truth-seeking group decisions.

Cue Utilization and Harvest Decisions in Commons Dilemmas

Donald W. Hine
University of New England, NSW, Australia

Robert Gifford and Yuko Heath
University of Victoria, BC, Canada
rgifford@uvic.ca

Ray Cooksey and Peter Quain
University of New England, NSW, Australia

This study examined how harvesters use feedback cues in a computer simulated commons dilemma, and whether or not lens model cue utilization patterns vary as function of harvesters' social and environmental values. 171 introductory psychology students from medium-sized universities in Canada and Australia managed a simulated fishery in groups of 3 over 60 seasons. Hierarchical linear modeling was used to

analyze the data. The Level 1 analysis revealed that participants took more fish during seasons when feedback (that is, cues) informed them that fish stocks, fish value, and fishing expenses were high, and when the noncooperative and cooperative group members (who were simulated by the computer) had taken more fish and fewer fish respectively during the previous season. The Level 2 analysis produced several cross-level interactions indicating that participants' cue use varied as function of their social and environmental values.

Brunswik's Methodological Fears are Music to our Ears

Patrik N. Juslin
Uppsala University, Uppsala, Sweden
patrik.juslin@psyk.uu.se

The music psychology group in Uppsala has been working on three projects with a distinctly Brunswikian flavor. In the project *Feedback-learning of musical expressivity (Feel-ME)*, we have developed a new computer program that automatically analyzes music performances to provide cognitive feedback (CFB) to performers in order to improve their communication of emotions. The program is based on the lens model adapted to communication of emotions in music performance (Juslin, 2000), and (in accordance with the original research on cognitive feedback by Hammond, Balzer, and others) offers the musician a chance to directly compare regression models of his or her playing strategies with optimal strategies based on regression models of listeners' cue utilization. The program allows the user to choose between outcome feedback and CFB indices, such as achievement, matching, consistency, and beta weights of both performer and listeners. The listener models are used to simulate listener judgments by means of *judgmental bootstrapping*. One of the biggest hurdles in the project was to develop computer algorithms that would automatically extract the relevant acoustic cues from music performances. In a recent study, 36 semi-professional jazz/rock guitar players were randomly assigned to one of three conditions: (1) CFB from the computer program, (2) feedback from music teachers, and (3) repetition without feedback (a contrast group). Performance measures revealed the largest improvement in achievement for CFB, though usability measures showed that certain aspects of the computer program could be improved (Juslin et al., 2004).

Last year, we published our review (Juslin & Laukka, 2003) of 104 studies of the nonverbal aspects of speech and 41 studies of music performance, which revealed parallels between the two channels concerning (a) the accuracy with which discrete emotions were communicated to listeners and (b) the emotion-specific patterns of acoustic cues used to communicate each emotion. The results could explain why music is perceived as expressive of emotion and are consistent with an evolutionary perspective on vocal expression of emotions. The discussion provided examples of how Brunswik's ideas could help to explain some of the peculiar results in previous research, for instance that successful communication was observed despite much inconsistency at the level of code usage. This thread was followed up, more extensively, in a recent book chapter on emotion in speech (i.e., Juslin & Scherer, in press), where we tried to demonstrate the benefits of Brunswik's paradigm for studies of nonverbal communication in general and emotional speech in particular. Specifically, we showed how Brunswik's insights might inform all sorts of methodological choices when designing a study in this domain.

Last but not least, we are currently awaiting a final decision about funding of a new project, *Appraisal in Music and Emotion (AMUSE)*. This project, which was largely inspired by my previous work on a book on music and emotion (Juslin & Sloboda, 2001), aims to develop a computational model of emotional reactions to music that combines different psychological mechanisms to explain and predict listeners' responses. We will probably put the lens model on ice for this project (but who knows). However, the project is still in line with Brunswik's approach (perhaps even more so than before) in the sense that we will try hard to investigate emotional responses to music as they spontaneously occur in everyday life, using electronic diaries and ambulatory physiological measurement to obtain representative samples of music experiences. This will perhaps be the most demanding project we have tried so far (if funded, that is), but this characteristic may also be in keeping with the nature of Brunswik's strivings.

References

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., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129, 770-814.

Juslin, P. N., Karlsson, J., Lindström, E., Friberg, A., & Schoonderwaldt, E. (2004). *Play it again with feeling: Feedback learning of musical expressivity*. Manuscript submitted for publication.

Juslin, P. N., & Scherer, K. R. (in press). Vocal expression of affect. In J. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *Handbook of nonverbal behavior research methods in the affective sciences*. New York: Oxford University Press.

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

Ecological Validity as a Mediator of Visual Attention Allocation

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

Last year, I used this opportunity to provide a broad overview of our work related to Brunswik. This year, I have chosen to take one particular piece of recent research and describe it in some detail, as I think it may represent a relatively new and interesting domain for applying the principles of probabilistic functionalism.

We have been conducting research for NASA evaluating the impact of Synthetic Vision System (SVS) displays for aviation cockpits. These displays use GPS and terrain mapping technologies to provide a (hopefully) veridical depiction of the out-the-window view even in low visibility (night, poor weather) conditions. A significant aspect of our work was a detailed analysis of pilots' eye movements in both SVS and non-SVS conditions, during the performance of an approach and landing scenario in a flight simulator. The experiment included both nominal and off-nominal conditions (e.g., an aircraft on the runway not shown by the SVS, the SVS image misaligned with reality, etc.)

While we obtained a number of interesting findings relevant to guiding the future design of such systems, our attempt to model the visual scanning patterns of pilots led to the project of interest here. How do you decide where to look when? Or when to look where? There is, of course, a large literature on visual attention

allocation. We decided to start with a model developed by my colleague Chris Wickens here at Illinois called the “**SEEV**” model. This model, informed by the results of scores of experiments, describes the probability of attending to a particular “Area of Interest” (AOI), among many such AOIs, as a weighted function of four attributes of the AOI: its **S**alience, the **E**xpectation one has that one will gain information from the AOI, the **E**ffort required to obtain that information (e.g., eye movements vs. head movements), and the **V**alue of the information gained to the task at hand.

In the formulation we originally took from Wickens’ research, the Expectancy (E) and Value (V) terms are combined multiplicatively, a choice motivated by utility maximization notions. In addition, Expectancy is measured as the bandwidth (rads/sec) of the information source, the idea being that AOIs with rapidly changing values will be sampled more frequently than more slowly changing AOIs. This assumption is grounded in the pioneering research of John Senders, who, in the mid 1960s, showed that humans’ visual scanning patterns will indeed become optimally adapted to signal bandwidth, as would be predicted by Nyquist’s sampling theorem. What we noticed in our review of this early research, however, was that Senders’ experiments, requiring participants to sample four gauges of varying bandwidths to detect alarms (dials out of range), perfectly confounded signal bandwidth (a highly salient proximal cue) and the true task criterion of alarm frequency. Thus, while Senders (and the SEEV model) assume that visual attention is adapted to bandwidth, in the light of our research on fallible SVS displays, we began to question this interpretation. Given the confound in Senders’ task (assuming a perfect correspondence between a display and its referent), we preferred to re-interpret Senders’ findings in accord with probabilistic functionalism: that his subjects had not actually adapted to bandwidth, but instead, had become adapted to the (in his case, 1:1) relation between the proximal cue of bandwidth and the distal task criterion of alarm frequency.

We searched the visual attention allocation literature without success to find studies conducted in probabilistic environments where the task involved the use of proximal cues with less than perfect correspondence to the task criterion (i.e., the literature assumes attention allocation is solely a function of the proximal environment). We thus constructed an experimental task modeled exactly after Senders’ own, but in addition to the perfect correspondence condition, we also included conditions where the various bandwidths of the four gauges had ecological

validities of 0.75 and 0.25 (correlations with alarm frequency).

Two central findings emerged. First, ecological validity had a pronounced effect in mediating both the rate of learning and visual scanning patterns, thus adding at least a fifth factor to the list of attributes of an AOI that mediate visual attention allocation. Second, we found that the best model, regardless of ecological validity, for predicting scanning patterns combined bandwidth and Value (which varied over gauges) in an additive, rather than multiplicative, fashion assumed in the SEEV model. Since the multiplicative formulation was largely created by appeal to theory (utility maximization) rather than data, Wickens and his colleagues thus re-analyzed their data from their own recent experiments and indeed also found that an additive formulation was superior to the multiplicative one, and have reformulated the SEEV model accordingly. We do not believe it is likely that our experimental design and findings would have emerged without an appreciation for probabilistic functionalism, the proximal-distal distinction, and task uncertainty. We are currently using the lens model to gain additional insights into our results, hopefully demonstrating its usefulness not only to the study of judgment, but also to the study of attention allocation as well.

Gender and Ethnic Background as Factors Influencing Personnel Selection Decisions

Pieter Koele and Els Ebels
Department of Psychology
Universiteit van Amsterdam
The Netherlands
p.koele@uva.nl

In a typical policy capturing study we investigated whether decisions made in personnel selection situations are affected by ethnic background and gender of the candidates. Participants (ranging from psychology students to lawyers) had to imagine they were a personnel manager at the Ministry of Justice, who had to select a (pre-screened) candidate for a position as policy co-ordinator in the Immigration Policy Department of the Ministry. For each of 64 fictitious candidates information was available on three explicit cues (the formal selection criteria): analytical skills (AS), oral fluency (OF) and writing ability (WA). Implicit cues were ethnic background (EB) and gender (G). Participants had to indicate for each candidate to what

extent they would feel inclined to invite him or her for an interview (INV).

METHOD

INV was measured on a scale ranging from 1 ("definitely no invitation") to 7 ("must absolutely be invited"). AS, OF and WA each could have ratings 7 (reasonable), 8 (good), 9 (very good) or 10 (excellent). By systematically combining the four possible ratings on each of the three cues the 64 vignettes were created. Ethnic Background was either native Dutch or immigrant Dutch, and Gender male or female. On the vignettes EB and G were introduced by a line which said: "Below you see the ratings on the three selection criteria of". In the blank space the name of the candidate was given. For example: "*Mr Klaas Voskuil*" would be a native male, and "*Mrs Fatima Boussaid*" an immigrant female. The immigrant names were Moroccan or Turkish (apart from the people of East and West Indian descent the largest immigrant populations in the Netherlands). To each of the four Gender/Ethnicity combinations 16 vignettes were randomly assigned. Statistical tests indicated that the four groups of vignettes did not differ significantly from each other on the three explicit cues. The judgment task, constructed in Authorware®, was administered to the participants on a computer screen. Responses were saved as an SPSS® data file.

ANALYSES

First, for each participant a regression analysis is done over all 64 vignettes, with INV as dependent variable and AS, OF, WA, EB and G as predictors, in order to describe the judgment process in terms of cognitive control (the multiple correlation Rs) and beta-weights.

In order to investigate the influence of EB and G on INV (the outcome of the judgment process) a 2 (EB) x 2 (G) ANOVA with repeated measures is done with INV as dependent variable.

Next, for each participant regression analyses are done over each of the four groups of 16 vignettes, again with INV as dependent variable, and this time AS, OF and WA as predictors. In order to investigate the effect of EB and G on the judgment process three 2 (EB) x 2 (G) ANOVAs with repeated measures are done with the beta-weights of AS, OF and WA as dependent variables.

RESULTS

A total of 74 participants took part in the experiment, 31 males and 43 females. Mean age was 27.6 years.

The regression analysis over all 64 vignettes demonstrates that participants had a high cognitive

control over their judgment strategies: Mean Rs is .86 (s.d. = .10). The beta-weights are .55 (AS), .46 (OF), .41 (WA), .02 (EB) and .00 (G). This result seems to suggest that the participants did not let their judgments be influenced by EB and G, but only by the three formal selection criteria.

However, the ANOVA on INV with EB and G as within factors yields a main effect of G and an interaction effect of EB and G. In general, females were more likely to be invited for an interview than males, and this effect is markedly stronger for natives (means of 5.05 and 4.35, respectively) than for immigrants (4.83 vs. 4.60).

In addition, the four separate regression analyses for each group of 16 vignettes also show a different picture than the one over all 64 vignettes. Cognitive control remains high (ranging from .84 to .92) but the beta-weights are clearly influenced by EB and/or G, as shown by the results of the ANOVAs.

On AV there is a significant main effect of Gender: the mean beta-weight is .49 for the male vignettes, and .61 for the female vignettes.

For OF we find a significant interaction effect of EB and G: mean beta-weights are .44 (native males), .51 (immigrant males), .49 (native females) and .42 (immigrant females).

And finally, on WA there is a huge main effect of Gender (mean beta-weight is .49 for male vignettes and .29 for female vignettes) and a somewhat smaller but still quite substantial main effect of EB: the mean beta-weight is .36 for native vignettes, and .47 for immigrant vignettes.

CONCLUSIONS

From a methodological point of view the results of the experiment are quite interesting. The overall regression analysis yields results that are quite misleading, in the sense that they suggest that ethnic background and gender of the candidates did not play a role in the judgment strategies of the participants. Only the more specific analysis for each of the four groups of vignettes separately indicate that participants followed to a large extent configural judgment strategies: the weights attached to the three formal selection criteria were clearly influenced by gender and ethnic background of the candidates.

From a psychological point of view the results are rather puzzling. At the moment we cannot find satisfactory explanations for the way gender and ethnic background influence the weights of the selection criteria. It is quite obvious that stereotyping plays an

important role, but in the absence of unequivocal theories or earlier findings on this specific topic it is difficult to formulate convincing interpretations of our empirical results. As usual, future research must try and find answers.

MCPL among children, adolescents and young adults: the case of inverse relationship

Peggy Lafon
Tours, France
lafon@univ-tours.fr

Few researches have been devoted to the study of the relation between age and MCPL. Some of them involved older people (Chasseigne et al., 1997, 1999, 2002, 2004), and very few included children and adolescents (Deffenbacher & Hamm, 1972; Miklich & Gillis, 1975; Montanelli, 1972). No MCPL study that involved children and adolescents implied learning of inverse relationships. Our intention was to provide data of such learning.

The present study was aimed at examining age differences in learning performance manifested among children, adolescents, and young adults placed in a multiple-cue ecology that was intended to be more complex than the very simple ones used in previous studies. The specific situation studied concerned a two-cue ecology. One cue was in direct relation with the criterion, as in previous studies, and the other cue was in inverse relation with the criterion.

Our overall hypothesis was that the learning of the inverse relation cue should find expression by utilizations of this cue more and more negative with age. This hypothesis was based on the different findings about cognitive development, specifically on the theoretical framework of executive functions (Zelazo et al., 1997). In view of the immaturity of executive functions among children (Kramer et al., 1994; West, 1996), it seemed that they would experience more difficulties, than adolescents and young adults, in learning of inverse relation cue. Specifically we expected that: (a) very few children under the age of 11 should be able to correctly utilize the inverse relation cue for predicting the criterion, (b) most adolescents and young adults over the age of 17 should be able to learn how to use the inverse relation cue, and (c) adolescents between 11 and 17 should show various levels of achievement.

A total of 439 participants aged from 5 to 27 years (191 males and 248 females) participated in this experiment. The materials consisted of three sets of 30 cards (10.5

x 7.5 cm), each showing two cue values in the form of vertical colored bars (blue). The cue values could take nine different heights, from 1cm to 9cm. For both cues, the distribution of the heights was approximately normal ($M = 5\text{cm}$, $SD = 1.95$). Between cues the correlation was non-existent. The criterion value was indicated at the back of each card and could take nine different numerical values, from 1 to 9. Its distribution was approximately normal ($M = 5\text{cm}$, $SD = 1.95$). The picture (21 cm x 29.7 cm) of a fruit juice machine was given to the participants. Two turning knobs appeared on the side of the machine. The correlation coefficients between the left-hand cue and the criterion was direct (.68), and the correlation coefficients between the right-hand cue and the criterion was negative (-.68). The level of uncertainty was $1 - .68^2 - .68^2 = .08$. The combination rule was additive.

The participants were shown the picture with the orange juice machine. They were told that the amount of liquid delivered by the machine could be controlled by the knobs. They were informed that their task was, (a) in a first session, to predict the amount of liquid from the settings of the knob(s), and (b) in four subsequent sessions, to learn to predict the amount of liquid by examining the back of the cards that bear the correct value. They were explained that the height of the bar(s) corresponded to the exact setting of the knobs on the machine. They were finally told that an exact prediction of the amount of juice was impossible because of a safety mechanism acting independently. The participants were shown five blocks of 30 trials. The experiment was self-paced, and the participants completed the task individually. They were instructed (a) to examine the graphical cue values on the card, (b) to predict the amount of liquid by giving a numerical value (1 to 9), and, except for the first block in which no feed-back was provided, (c) to turn over the card and consider the actual amount also indicated as a numerical value (1 to 9). The experiment tasks took from 30 to 40 minutes to complete.

Our hypothesis were well supported by the data. As expected, the younger children always used the inverse relation cue in a direct way; the others progressively ceased to use the inverse relation cue (mostly the 9-17 year-olds). A substantial negative utilization of the inverse relation cue was not observed before the age of 11. Only 23% of the 11-17 year-olds appeared to be able to restructure the representation they have of the task; that is, to select the inverse relation hypothesis during learning. It was only among the young adults that we could observe a majority of participants using the inverse relation cue in a correct way. These results are interpretable within the

theoretical framework offered by the “executive function” construct (Zelazo et al., 1997). In this situation involving both inhibiting a prepotent response and coordinating two cue values of opposite meaning (e.g., direct relation and inverse relation); that is, a situation involving a higher level of executive functioning, most young children, many adolescents and some young adults tested in the present study failed.

Lafon, P., Chasseigne, G., & Mullet, E. (2004). Functional Learning Among Children, Adolescents and Young Adults. *Journal of Experimental Child Psychology*, 88, 4, 334-347.

Metacognition and Value Attribution

Guillermo Macbeth
Universidad del Salvador, Argentina
macbeth@fibertel.com.ar

The last year I concluded my doctoral thesis under the supervision of Prof. Nuria Cortada de Kohan in Universidad del Salvador, Argentina. The focus of this research was on the metacognitive monitoring of value attribution processes. The experimental task required the introspective evaluation of some decisions concerning the attribution of truth, falseness, or novelty value to some previously seen statements. The brunswikian key of the study was the identification of an adaptive metacognitive heuristic. The priority order in the subjective selection of the source of information resulted convergent with the priority order of accurate attributions. These outcomes suggest that the metacognitive monitoring of value attribution processes could be considered efficacious, i.e. fallible but adaptive. The general heuristic activated spontaneously in this experimental task probably defines a pattern of preferences, i.e. subjects try first to remember clearly the corresponding value, if that does not work they search some fuzzy cue like a cognitive sensation (e.g. familiarity) or conjecture. If this second source is not available, then subjects simply guess. This heuristic gets spontaneously activated and their results could be considered as very satisfactory to achieve environmental adjustment.

Extending this interest on metacognition and decision making I am now planning some experiments to study the monitoring processes of the overconfidence effect under the supervision of Prof. Rudolf Groner from the University of Bern, Switzerland.

Brunswik and Human Factors Research

Kathleen L. Mosier, Ph.D.
Industrial/Organizational Psychology
San Francisco State University
kmosier@sfsu.edu

In my research lab at SFSU, we are getting ready to begin research investigating the impact of several contextual variables, such as presentation mode (of cues and information), risk, accountability, and environment (electronic vs. naturalistic) on decision strategy (coherence or correspondence). Hopefully, I will have something to tell you about this project next year!

One accomplishment for this year was the acceptance of a Brunswik-focused panel for the Annual Meeting of the Human Factors and Ergonomics Society - BRUNSWIK'S LENS MODEL IN HUMAN FACTORS RESEARCH: MODERN APPLICATIONS OF A CLASSIC THEORY. I worked with Alex Kirlik to create a session that presented the basic tenets of the lens model, and discussed its viability and applicability in the context of modern human factors investigations, and in particular, for explorations of human judgment in high-technology environments. I include a synopsis of my organizational framework and the topics of each of the panelists below – if you would like to contact any of our panelists, please let me know and I will arrange it! The session was well attended and well received – so I hope that we can look forward to more work in this venue among the human factors community.

Recently, some human factors researchers have utilized the Brunswikian framework and the Lens Model to perform judgment analyses in high-technology environments. Technology creates a new human-task environment ecology, and is both a facet of the judgment system, and a mediator within it. Because of technology's multiple roles, the Brunswikian framework can be utilized in several overlapping configurations – to compare, for example, judgments of *technology as decision maker* with those of human decision makers in a task environment; to explore human judgment with *technology as cue* within an array of cues; or to capture judgment strategies with *technology as environment*, or compare them with judgments in non-technological environments. For this panel, Ann Bisantz began by providing an introduction to the Brunswikian viewpoint, and discussed its potential for examining human-machine systems applications. Amy Pritchett discussed facets of the judgment task in combined human-machine systems, and usage of the lens Model judgment analysis to study these tasks. Ellen Bass

presented HAJL (Human-automated Judgment Learning), a methodology for exploring the interaction between human and automated judges. Ling Rothrock discussed limitations of the standard Lens Model Equation, in terms of heavy reliance on regression analysis, and outlined the possibility of extending the model through the integration of rule-based models of judgment. Lastly, Alex Kirlik placed the concept of representative design into historical perspective, and established a link between it and human factors requirements as described by Alphonse Chapanis, a human factors pioneer.

Payoff Perception in Negotiations

Jeryl Mumpower
University at Albany, SUNY, USA
mumpower@albany.edu

This past year, I once again spent most of my time doing administrative tasks, because I am now Interim Provost of my university. I reported last year about a paper coauthored with Tom Stewart and entitled "Detection and Selection Decisions in the Practice of Screening Mammography." This paper was published this year in the Journal of Policy Analysis and Management.

Also published this year in Group Decision and Negotiation, after a gestation period of about 9 years, was a paper co-authored with Jim Sheffield, University of Auckland, Tom Darling, University of Baltimore, and Rick Milter, Ohio University and entitled "The Accuracy of Post-Negotiation Estimates of the Other Negotiator's Payoff". This paper reported two empirical studies of interpersonal understanding in negotiations. In the first study, the accuracy of post-negotiation estimates of the other negotiator's payoff was assessed after a role-playing simulation. Only a minority of participants exhibited evidence of the fixed pie bias, in which negotiators view all negotiations as distributive, fixed-sum situations.

Participants' estimates of the other negotiator's payoff were generally better fit by the equal payoffs model, which presumes that participants believe the other negotiator's payoff is the same as one's own. This held true for both distributive task structures in which the fixed-pie view is descriptively appropriate and integrative negotiation task structures in which the fixed-pie view is inaccurate. The results did not support the hypothesis that superior understanding about the other negotiator's interests helps negotiators

to achieve better outcomes for themselves; the correlation between predictive accuracy and the value of participants' own payoffs was generally low.

A second study was conducted to test the hypothesis that negotiators typically see negotiations as fundamentally a distributive, fixed pie problem, but believe their own negotiated agreements yield roughly equal payoffs to both negotiators. The results supported this hypothesis. In this second study, participants estimated the other negotiator's payoffs over a sample of hypothetical contracts. The payoff schedule estimation procedure, which has been widely used in previous research, was not used in the present research because it was shown (through Brunswikian-inspired analyses) to have serious methodological, conceptual, and procedural shortcomings.

Appropriate Judgment and Appropriate Medical Treatment Decisions

Roy M. Poses MD
Brown University Center for Primary Care and Prevention, USA
Roy_Poses@brown.edu

With Jamie Brehaut, PhD, and Jeremy Grimshaw MD of the Ottawa Health Research Institute and the University of Ottawa, Canada, I am working on another project to try to understand why physicians fail to make treatment decisions despite the best available evidence from clinical research, as recommended by proponents of evidence-based medicine. We have submitted a grant application to address why physicians often prescribe antibiotics for patients with sore throats who do not clearly have streptococcal infections ("strep throat"). In the absence of such infections, antibiotics do no demonstrable good, may still cause side-effects, and their widespread use may increase the number of resistant micro-organisms. We also propose to address why physicians fail to prescribe cholesterol lowering drugs ("statins") for patients who have coronary artery disease, despite strong evidence that prescribing these drugs may prolong life and prevent further cardiac problems for some patients. We hypothesize that physicians over-estimate the benefits and under-estimate the harms of antibiotics in the first case, and under-estimate the benefits and over-estimate the harms of statins in the second case. Furthermore, we plan to use the Brunswik lens model to determine if such poorly calibrated probabilistic judgments arise from using non-predictive cues, or

failing to use or mis-weighting predictive cues as compared to models of the ecology. The grant application is currently under review by the Canadian Institutes of Health Research.

Solving the "User Override Problem": Why Do Experts Turn Off Decision Support Systems?

James Shanteau
Psychology, Kansas State University, USA
shanteau@ksu.edu

David Gustafson
Computer Science, Kansas State University, USA

INTRODUCTION: In the climactic scene in the original Star Wars movie, Luke Skywalker is the last hope of the Rebels in their crucial battle against the Death Star. "Just as Luke is about to be eliminated by Darth Vader, he hears Obi's familiar voice reassuring him: 'Use the Force, Luke. Let go, Luke. Luke, trust me.' Relying on the Force Luke switches off his targeting computer, using his own intuition instead" (from the Official Star Wars Script Synopsis, Dirks, 1990).

The lessons from this fictional episode reflect several widely-believed truths: (1) Good triumphs over evil. (2) There are greater forces inside us than we know. (3) The key is to turn off the computer!

In an informal survey we conducted, virtually everyone contacted understood precisely why Luke turned off his targeting computer. It was seen as a natural, even inevitable, course of action. The question we are asking is: Why is intuition seen as superior, not just to formal human analysis, but to optimally-designed computer systems?

BACKGROUND: Decision Support Systems (DSS) have been developed to assist decision makers in a variety of technologically complex tasks. Typically, human experts provide some or all of the inputs and the system will compute an optimal course of action. Such systems have been developed for medicine, aviation, geology, military intelligence, accounting, computer networks, and space flight (to name a few).

Despite the sophistication of DSS, there are many anecdotal accounts of users overriding, and in even turning off, the system in critical situations. In air traffic control, for instance, controllers ignore aircraft flight-path forecasting tools in high traffic conditions, mainly because of cognitive overload.

The irony is that these are precisely the types of high-consequence situations that such systems were designed for. Thus, despite the enormous investment in developing DSS, experts frequently turn them off (or override them) in crucial settings. At the same time, experts often use these same systems in everyday, straightforward situations where they need less help.

Two questions behind the present research program are: (1) why do experts mistrust DSS at critical times? And (2) what can be done to improve design of such systems to address this mistrust?

We are addressing these questions in a three-phase research program. The first phase is to identify the psychological and technical factors behind the "User Override Problem." The second phase is to investigate how these factors impact decisions by experts to turn off DSS. The final phase will be to propose and evaluate new software design strategies to address the reasons why experts ignore DSS.

PSYCHOLOGICAL FACTORS: In this report, we will describe some of the psychological factors we have found that are involved in decisions to override DSS. They are listed alphabetically.

(1) **Attention Overload:** It is a well-known psychological phenomenon that information overload can impair performance. Experts have a unique ability to simplify complex problems and, as a result, to avoid information overload. They have little tolerance for irrelevant or even partially relevant information sources. Even the best-designed DSS requires some attention from experts, eg, through alarms or flashing lights. In critical settings, these warning signals can be an irritating distraction (a form of irrelevant information) and thus are ignored by experts.

(2) **Conflict of Mental Models:** The mental model/problem representation is vital to the decisions made by the expert/DSS. In normal situations, the expert's mental model and the system's representation usually agree. In more difficult situations, however, there can be a conflict between the mental model of the expert and the problem representation built into the DSS. This "mode awareness mismatch" can seriously compromise the quality of the DSS output and result in the expert turning off the system.

(3) **Subjective Cost/Benefit Analysis:** In most decision situations, the expert user must decide whether it is worth the time and trouble to use the DSS. In effect, the expert performs a subjective cost/benefit analysis to decide whether the cost of using a DSS is too great given the limited benefits there may result from a

computerized decision tool. Often, the expert will determine that a quick, approximate answer is preferable to an optimal, but slower answer.

(4) Designing to the Average: Computer scientists design decision support systems to maximize the success rates for typical (or average) cases. By definition, these are not the tough, but rare cases that challenge experts and lead them to look for assistance. Therefore, there can be a disparity between the situations that DSS are designed to handle and the situations where really need them. It should not be surprising, therefore, that experts may ignore DSS in the toughest cases.

(5) Legal Implications: Most DSS are based on well-specified operational rules that provide a precise reconstruction of how a decision was made. That is, they are transparent to users – and to lawyers. This opens the door for lawsuits and legal challenges should the decision outcome go wrong. Fear of lawsuits has already led to reports of the purposeful “dumbing down” of certain guidance systems. Therefore, experts may be reluctant to rely on DSS because of the purposeful reduction of capability.

(6) Miscalibration/Recalibration: Based on expert inputs, intelligent systems usually yield more extreme outputs than comparable human decisions. When experts see the outputs from such systems, they often recalibrate their inputs to produce a more desired output. For routine cases, these adjustments usually work smoothly. For more unusual cases, it may be difficult, if not impossible, to recalibrate, leading the expert to ignore the system in tough situations.

(7) Perceived Risk: Incorporating a DSS into the decision loop almost always leads to an increase in the subjective complexity of the process. As complexity increases, perceived risk typically increases. Since most humans (including experts) are risk adverse, they will take steps to reduce risk, especially for important or consequential decisions. For experts, these steps can include ignoring or turning off the DSS.

ONGOING RESEARCH: We are now investigating the relative importance of these and other factors in experts decisions to use or not to use DSS in advanced guidance systems (such as air traffic control). Our goal is to identify the most important factors by a using a combination of interviews with experts and experimental research to estimate cue weights.

Our next step will be to establish new directions for improving the design and evaluation of DSS software. Our goal is to establish guidelines for best practices to

be used in writing new decision support systems. If miscalibration of inputs is a concern, for example, then routines for encouraging accurate calibration will be developed.

For further information on this project, contact James Shanteau at shanteau@ksu.edu.

Nursing Attitudes and Clinical Inferences in Psychiatry

Lars Sjö Dahl
Lund, Sweden
le.sjodahl@swipnet.se

Great concern has been expressed in the literature about the extent to which the latest criteria manual DSM-IV allows false, positive inferences (Spitzer, Wakefield, 1999). This problem is difficult to tackle because lack of appropriate validity criteria. The concept ecological validity might be useful in this context. Ecological validity was introduced as a psychological term by E. Brunswik (1952) to designate the strength between a perceptual cue and an object in the person's environment. Today ecological validity has acquired a wider meaning referring to generalizability of observations and inferences to contexts outside the observation situation. This was however not the original meaning of the concept (Hammond 1998).

Criteria for ecological validation of clinical inferences in psychiatry (mental health field) are to be sought for in the patient's daily life outside institutions. Thus attention has to be directed to the patient's ecological niches or idiographic fields like family-life, work environment, leisure time activities, primary-groups, relatives etc. The diagnostic dialogue between doctor (psychologist) and the patient is the steppingstone for the inferences written in the patient chart, inferences that ought to be validated against idiographic information communicated by the patient in cooperation with an independent researcher, retrospectively evaluating the content in the patient chart. This research approach, cooperation with the patient, has been emphasized several times in the medical literature. In *Schizophrenia Bulletin* (1995, 21(3), pp. 279-284) we find the following urgent request: “----We welcome other contributions from patients, ex-patients, or family membersClinicians who see articulate patients with experiences they believe should be shared, might encourage these to submit their articles to---“. In *The Lancet* (2000, 355:1540-1543)

Herxheimer et al present the database called Individual Patients' Experiences of Illnesses.

Further, there might be a special reason to retrospectively complement and re-evaluate diagnostic inferences in psychiatry by means of patient's first person accounts. The reason is that it is often taken for granted that the doctor and the patient share a common approach (cooperate) in their search for valid diagnostic inferences (coming to terms), that would suggest the most appropriate treatment. This assumption can, however, be completely wrong. The patient seems often to contribute to the diagnosis in an inductive way, hoping that descriptions of symptoms would help the doctor to arrive at an aetiological understanding (diagnosis) of an underlying illness that might explain the symptoms. The doctor, however, might have the opposite approach, working deductively from nomothetical, nosological broad concepts, trying to justify his perhaps premature "labeling" of the patient by selecting "suitable" cues among the patient's many symptom descriptions. How do such opposite approaches in a situation, demanding cooperation between two persons, affect reliability and validity of those inferences that are written in the doctor's notes, information that will be the basic ground for choosing what could be considered as the best treatment in the single case? This choice implies careful weighing of treatment benefits against risks for adverse side-effects that in many cases result in irreversible, severe iatrogenic suffering for the patient.

My research in patient-caregiver relations spans a rather wide area, some examples:

1. Content analysis of curriculum for nursing education
2. Content analysis of an attitude-item pool
3. Construction of attitude scales, estimating nurses attitudes to the expressive (need-related) aspect of patient care.
4. Mapping attitude and personality changes during the 5-term education for nurses, cross-sectional study
5. Factor analyses of nursing attitudes
6. Work analysis with critical incident interviews confined to the expressive (need-related) aspects of nursing care.
7. Retrospective analysis of 20- patient charts combined with patients' first person accounts resulting in a category system (taxonomy of error) for false positive, diagnostic inferences (34 categories).

8. A preliminary suggestion for idiographic fields (ecological niches) is presented, which might be of some help in searching for relevant cues in the patient's environment.

9. Interview-guides and questionnaires for mapping idiographic fields and cues are outlined

Do Physicians Recalibrate Patients' Pain for Functional Reasons?

Paul Sorum

Albany Medical College, Albany, NY, USA

PaulSorum@cs.com

If you permit me to stretch the definition of Brunswikian, my recent collaboration with Laetitia Marquié, a PhD student in psychology at the University of Toulouse, and with her advisers, Eric Raufaste, Dominique Lauque, and Claudette Mariné, might squeeze in.

In order to explore her previous finding that emergency department physicians gave systematically lower ratings to patients' pain, both on arrival and at departure, than did the patients themselves, Laetitia presented 52 ED physicians with 45 scenarios of 30-year-old males presenting with abdominal pain and asked them to rate the patients' pain (on a visual analogue scale of 0-10). These scenarios consisted of all combinations of 3 types of cues: 1) 3 configurations of highly-correlated signs of the severity of the abdominal pathology (based on height of temperature, degree of abdominal guarding, and number of polymorphonuclear leukocytes in the blood); 2) 5 levels of the patients own pain rating (1, 3, 5, 7, or 9); and 3) 3 levels of highly-correlated physical manifestations of pain (involving degree of moaning and of facial expressions of pain).

As reported in a poster at the recent meeting of the Society for Medical Decision Making, regression analysis found that the mean beta weights of the 3 cues were signs of abdominal pathology 0.15, patient's pain rating 0.60, and manifestations of pain 0.49. When the mean physicians ratings were plotted against the patient ratings in the scenarios, the physicians recalibrated the patients' ratings increasingly upward as the patients' ratings descended below 5 and increasingly downward as they rose above 5. This was, of course, a reflection of an orthogonal rather than a representative design. I calculated a "discordance" score for each scenario—the degree to which the other 2 cues were discordant with the patient's pain rating—and we decomposed each point on the physician's rating curve (for each patient rating) according to the

degree of discordance from the particular patient rating. Even though, for simplicity, I gave equal weights to abdominal pathology and manifestations of pain (in spite of the differences in the beta weights), the result was a wonder to behold (at least for me). When the discordance was 0, the physicians gave almost the same ratings as did the patients. The decomposed points gave a series of nearly straight, parallel lines; this demonstrated that, for each level of patient rating, the physicians (on average) recalibrated the patient's pain in a consistent manner upwards or downwards in response to the degree and direction of discordance.

We concluded that, just as previous studies by us and others have shown that physicians readjust patients' pain ratings in response to "non-functional" factors such as the age, sex, and ethnicity of both patients and physicians, physicians also make use of more legitimate "functional" factors (how much pain the patient's illness should cause and how much pain the patient is manifesting). Laetitia also asked the physicians about treating the patients' pain; I will learn the results when I read her recently-completed thesis.

Collaboration in NMCPL

Maarten Speekenbrink
Psychological Methodology,
University of Amsterdam, The Netherlands
m.speekenbrink@uva.nl

As part of my dissertation, I have been working on group collaboration in Nonmetric Multiple Cue Probability Learning (NMCPL). In classical research on interpersonal learning (IPL), participants are trained to infer the value of a criterion on the basis of a single cue (two cues are provided, but only one has predictive validity). In the group task, the ecological system is changed, so that multiple cues are related to the criterion. A problem with this paradigm is that it is difficult to distinguish the effect of interpersonal learning from that of changing the ecological system. Indeed, as later research by Andersson and Brehmer (1979) showed, the earlier results could be mainly attributed to the latter effect. To avoid this problem in the experiments I conducted, the ecological systems in which participants were individually trained were subsets of the complete ecological system provided in the group task. Interpersonal learning was thus learning about the relation between new cues and the environment. This experimental setup makes it possible to study group performance under different distributions of information over the group members.

On the methodological side, Björkman's (1973) approach for the analysis of individual behavior in NMCPL tasks was extended to ecological systems with more than one cue. I have derived new cue validity and utilization coefficients for NMCPL based on information theory. Information-theory provides a very promising framework, especially for NMCPL, but also for MCPL, and I'm planning to do more work in this area.

For collective behavior in NMCPL specifically, I have investigated optimal and sub-optimal group processes. More specifically, I investigated the adequacy of three processes: simple majority, weighting-by-achievement and weighting-by-evidence. The achievement associated with these processes depends crucially on how information is distributed over group members. The first experiment investigated differences in group achievement between groups in which all information was shared and groups in which the information was completely distributed. In the second experiment, group achievement was studied when information was partly shared, partly unique, with different levels of cue validity of the shared cue. Overall, groups appeared to work under a weighting-by-confidence process. Since confidence is related to achievement, this results indirectly in a weighting-by-achievement process. Furthermore, group achievement was higher in the condition where the information was completely distributed, than in the condition in which the information was completely shared.

References

- Andersson, H., & Brehmer, B. (1979). Note on the policies acquired in interpersonal learning. *Organizational Behavior and Human Performance*, 24, 195–201.
- Björkman, M. (1973). Inference behavior in nonmetric ecologies. In L. Rappoport & D. A. Summers (Eds.), *Human judgment and social interaction* (pp. 144–168). New York, NY: Holt, Rinehart and Winston.

Quantity Judgment with Conflicting Information

Michael Tschirhart
Department of Psychology
University of Michigan, USA
mtschirh@umich.edu

J. Frank Yates
Department of Psychology and Ross School of
Business, University of Michigan, USA
jfyates@umich.edu

Our current research concerns how people make quantity judgments of a criterion based on a set of predictive cues that contains conflicting information. The focus of this work is on how the presence of those conflicting cues influences (1) the accuracy of quantity judgments, (2) confidence in those judgments, and (3) the processes that generate those judgments.

We have conducted simulations which show that (1) the prevalence of conflicting cue cases in common judgment situations is considerably greater than intuition might suggest, and (2) the overall accuracy of judgments derived from leading behavioral models in such situations can vary substantially.

Using data from an important real-world ecology (student performance, indexed in a national education study), a combination of simulation and empirical methods are being applied to critically evaluate competing accounts of the processes (e.g., discounting and similarity) giving rise to point and confidence judgments in the presence of conflicting cues. Besides providing representations of those processes, our simulations are intended to provide practical insights on the accuracy of the alternatives.

A brief report on a few somewhat Brunswikian activities, so called because they are concerned with complex, probabilistic worlds.

Alex Wearing
University of Melbourne, Australia
ajwe@unimelb.edu.au

The complex decision group (Mary Omodei, Jim McLennan & Alexander Wearing) has been studying dynamic decision making for a number of years using a combination of field research and computer simulations realised as a fire fighting task. The choice of cover story is having a happy consequence as the Australian Government has just begun funding a multi-million dollar study of various aspects of bushfires (wildfires to some of you), one area of which will involve decision making on the fire ground with a particular concern with

safety i.e. making risky decisions where a miscalculation could lead to incineration. This work is just starting, so we have only intriguing lines of inquiry to pursue at the moment.

Associated with it is a study of volunteerism, which involves a decision to join the fire service as a volunteer, and then a subsequent decisions to remain or leave, a decision sequence which may be repeated more than once. This decision is one with social consequences, as rural fire services in Australia are mainly staffed by volunteers.

Meanwhile, our laboratory work continues, with our two aspects of our current focus being on meta-cognition and situational awareness, which we are casting in terms of issues of the economics of cognition. These terms have become popular, and we are currently looking for agreed definitions of their meaning. We are also interested in good ways to measure them. We have been looking at the human factors aspects of network-centric warfare (NCW), an idea whose time has come, and may even be gone, or a least going. The idea is that, backed by modern communications technology, the fog of war is a phenomenon of the past. We have run a number of experiments that show that the glare of war is just as problematic as its fog.

A third aspect of our current focus is multiple person decision making, an issue of some importance in a number of practical situations, where individuals have to work together as a team. To date, most research into this topic has been concerned with team management of small crews, e.g. flying aircraft, who employ highly skilled, but nevertheless somewhat routine procedures. Our focus is on situations where somewhat ad hoc teams comprising members of varied training and experience must come together e.g. to fight a forest fire, where adaptation to the demands of both the task environment and the heterogeneity of team composition is problematic and requires effective resource allocation-material, individual and team. At this point we are still at the stage of planning to collect field data during the Australian summer fire season.

We also have a website to which the curious can repair: www.latrobe.edu.au/psy/research/cdrj. It will tell you what we have been and are doing

Travelling at an Interface

Elise Weaver

Worcester Polytechnic Institute, USA

eweaver@wpi.edu

This has been an unusual year for me, with opportunities to travel because of my idiosyncratic niche at the interface of judgment and system dynamics.

One interesting opportunity came when I was invited to be "first opponent" at a Norwegian dissertation defense by Agata Sawicka regarding her system dynamics approach to decays in security compliance.

My work with Tom Stewart continues, with papers under revision regarding the dimensions of judgment tasks, and the commonalities between implicit learning and vicarious functioning. I am also working with Jim Holzworth on the adaptation of my judgment analysis task software to the needs of his graduate student, Amy Reese, who is interested in investigating Hammond's Cognitive Continuum Theory.

Work that I did with George Richardson putting Ken Hammond's (1996) Taylor-Russell diagram discussion into a system dynamics form has interested people outside psychology. The interest comes from a pressing need for a standardized indicator of building security and a way of deciding the appropriate security threshold within the context of complex social and economic trade-offs. I now have an "ambassador" talk to non-Brunswikians and non-system dynamicists about three possible tools to address these issues: 1) judgment analysis, 2) the Taylor-Russell diagram, and 3) system dynamics. A colleague affiliated with Fire Protection Engineering at WPI, Brian Meacham, invited me to speak on this issue at an international conference in Washington, D.C. regarding the establishment of performance-based building codes. In addition, a colleague of his, Richard Little of the National Research Council, invited me to speak at a workshop including General Accounting Office representatives on the same topic at separate conference.

On the theoretical side, I am continuing with my pet project of a geometric approach to judgment analysis that I presented last year at the Brunswick meeting. On the applied side, I find that I am more and more interested in issues relating to public participation in democracy (perhaps because I am an expatriate Canadian in the U.S. during an election year).

Policy Recognition in Practitioners' Antibiotic Prescription

Bob Wigton

University of Nebraska Medical Center, USA

wigton@unmc.edu

As part of a larger project attempting to reduce antibiotic use in respiratory tract infection, Carol Darr and I have been studying community practitioners' use of clinical and patient cues in deciding whether to prescribe antibiotics. The current study was inspired by Reilly and Doherty's work showing that some decision makers could recognize their own policies. Since we had just completed policy profiles on community practitioners, we decided to test whether they could recognize theirs.

Because of the number of participants (54), we clustered the individual policies to obtain 9 clusters of related policies. We selected the policy of the cluster member that was closest to the average weighting in the cluster as an *archetype* for the cluster. These cue weights of these 9 policies were shown to the practitioners as histograms. Their own policy was substituted for the archetype of their own cluster. What we found was that the practitioners did no better than chance at choosing their own policy from among the 9.

Currently, we are interviewing each of the practitioners to see if we can understand why they were unable to identify their policy. Both the type of decision and the methods differed from Reilly's so we're planning some further studies to explain the difference in outcome.

Exploring the Continuum between Compensatory and Noncompensatory Decision Strategies

Jing Yin

jxy178@psu.edu

Ling Rothrock

lrothroc@psu.edu

Pennsylvania State University, USA

We are currently engaged in research to extend the Genetics-Based Policy Capturing (GBPC) method (Rothrock and Kirlik, 2003) to formulate a complement, rule-based model, to the traditional lens model formulation. We propose a rule-based formulation of the lens model based on the notion of percentage match correspondence, which could account for noncompensatory decision strategies under time-stressed and information-rich environments. The psychological interpretation of lens model parameters

under the proposed model is redefined so that G is designed to estimate how well the rule-based environment model corresponds to the noncompensatory policy-capturing model of the human decisions. Moreover, C will indicate the extent of knowledge that is not captured by the rule-based modeling.

Encouraged by research findings which suggest that humans tend to change decision strategies based on the amount of information demands (Payne, 1976) and time available (Wright, 1974), we speculate that there exists a continuum between compensatory decision strategies at one extreme and noncompensatory strategies at the other. Furthermore, we hypothesize that this continuum can be described via cue-criteria relationships for both the ecology and the human cognitive system. A simulation study has been conducted to investigate techniques to capture shifts along the continuum.

In our simulation, judgment and criterion instances are created by mixing the data generated by specified linear-additive-based and rule-based generators. Furthermore, all the instances are ordered along a scale which represents the proportion of linearly-generated judgments and criteria. For each instance, both compensatory and noncompensatory lens model analyses are applied to calculate two sets of lens model parameters. Using a computer simulation, we sought to detect a relationship between the location of a point on the proportion scale and the location of a point on the compensatory-noncompensatory strategy continuum.

Results from the simulation suggest that C values under noncompensatory lens model analysis is similar under different parameter settings. The higher the proportion of linearly generated data, the higher C becomes. By further analyzing the data, we ran regression analyses to test if the relationship between C and the portion of linearly generated data (which we call P) holds. Results show that a perfect goodness of fit exists when we use the regression line to depict the relationship between C and P (high R^2 around 1).

While further investigation with actual human performance data is needed, our preliminary results are encouraging. We hypothesize that if we know when compensatory and noncompensatory strategies tend to be used to cope with different environments, we then can predict the location of the decision strategy along the compensatory-noncompensatory continuum.

Payne, J. W. (1976). Task Complexity and Contingent Processing in Decision Making: An Information

Search and Protocol Analysis. *Organizational Behavior and Human Performance*, 16, 366-387.

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

Wright, P. (1974). The Harassed Decision Maker: Time Pressures, Distractions, and the Use of Evidence. *Journal of Applied Psychology*, 59(5), 555-561.

Tentative Agenda

*The 20th Annual International Meeting of the Brunswik Society
Minneapolis, MN
November 18-19, 2004
Millennium Hotel, Avenue Room 2 & 3*

Thursday, 18 November

12:30 - 13.00 **Late Registration**

13.00 – 13.20 **Introductions and Welcome** (Jim Holzworth, Elise Weaver, Tom Stewart)

13.20 – 14:20 **Paper Session 1: The Subject in the Ecology** (Chair: Elise Weaver)

Asaf Degani, Mike Shafto, & Alex Kirlik
What makes vicarious functioning work? Exploring the geometry of human-automation interaction

Rob Youmans
A case for cognitive information feedback's utility in certain judgment situations

14:20 – 15:20 **Paper Session 2: The Task Ecology** (Chair: Jim Holzworth)

Jörg Rieskamp & Anja Dieckmann
Information redundancy influencing probabilistic inferences

Mike Doherty & Richard Anderson
Analyses of statistical environments from which inferences are drawn

15.20 – 15:40 **Tea and Coffee Break**

15.40 – 16:40 **Paper Session 3: Innovations in Modeling** (Chair: Bernhard Wolf)

Tom Stewart, Rob Hamm, & Tom Tape
A new formula and illustrative data for the logistic form of the lens model equation

Rob Hamm
How much structure is appropriate for a judgment model?

16:40 **Adjourn**

18:30 – Evening **Group Dinner at a Restaurant in Minneapolis**

Friday, 19 November

08.30 – 09.00 **Tea and Coffee Break**

9:00 – 10:30 **Paper Session 4: Applications in Education, Law and Foreign Policy**
(Chair: Jim Holzworth)

Bernhard Wolf
Representative longitudinal design and persistence - empirical results in educational psychology

Marcio Carvalho
The use of judgment analysis and social judgment theory in legislative research

Phil Dunwoody & Ken Hammond
The foreign policy of preemption via the Taylor-Russell diagram

10:30 – 10:50 **Tea and Coffee Break**

10:50 – 12:00 **Symposium** (Chair: Jeryl Mumpower)

Could The Brunswick Society Fix The Broken Parts of the U.S. Intelligence System?

Panel: Ken Hammond, Elise Weaver, Phil Dunwoody

12:00 – 13:30 **Buffet lunch** followed by an invited presentation by Bob Bateman, Certified EASI Tennis® Professional

13:30 – 14:30 **Paper Session 5: Training and Expertise in Professional Judgment**
(Chair: Phil Dunwoody)

James Shanteau, Brian Friel, Rick Thomas, & John Raacke
Assessing training in an air traffic control microworld

David Weiss, Jim Shanteau, & Priscilla Harries
Objective analysis of professional opinion

14:30 – 15:30 **Paper Session 6: New Applications** (Chair: Tom Stewart)

Konstantinos V. Katsikopoulos, John M. C. Hutchinson, Peter M. Todd, & Bartosz Gula
Do automobile drivers optimize or satisfice when they park?

K. Balaji Rao, M.B. Anoop, N. Lakshmanan, S. Gopalakrishnan & T.V.S.R. Appa Rao
Application of Brunswikian theory for corrosion damage assessment of reinforced concrete structural members

15:30 – 15:50 **Tea and Coffee Break**

15:50 – 17:20 **Paper Session 7: Applications in Medical Judgment and Decision Making** (Chair: Neal Dawson)

Bob Wigton, C.A. Darr, B. Leeman-Castillo, & R. Gonzales
Practitioners' insights into their own decisions about prescribing antibiotics in respiratory infections

Lars Sjö Dahl
Nursing attitudes towards patients' psychological needs

Christine Huttin
New cost sensitivity measures to assess impact of reimbursement systems on physicians' treatment choices, based on an adaptation of the lens model on cost related cues

17:20 – 17:30 **Brunswick-Hammond New Investigator Prize** (Awarded by Ken Hammond)

17:30 – Farewell and Meeting Adjourned