THE BRUNSWIK SOCIETY Newsletter 2000

Volume 16, Millennium Issue
Albany, New York
Fall, 2000

This edition of the Brunswik Society Newsletter was edited by Elise Axelrad Weaver, with the assistance of Suzanne Wissel (swissel@csc.albany.edu),and supported by the Center for Policy Research (University of Albany, SUNY) and the Center for Research on Judgment and Policy (University of Colorado, Boulder). The HTML version was prepared by Junseop Shim. The Newsletter was reformatted as a pdf file in 2022, and some minor corrections were made.

Contents

Otitis, Multiple Cue Probability Learning, and Control Theory

Claudia Gonzalez Vallejo Athens, OH

Optimal Hierarchical Command Styles in Dynamic Decision Tasks

Alexander Wearing Melbourne, Australia

Uncertainty, Confidence and the Lens Model

Ann Bisantz Buffalo, NY

Self-regulation, Feedback and Leadership Skills

Shawn Noble Boston, MA

The Use of Configural Information in Multi-Cue Judgments by Individuals and Groups

Scott Tindale Chicago, IL

Dealing with Conflicting Information

Kathleen L. Mosier, Jeffrey Keyes, and Roberta Bernhard San Francisco, CA

Integration of Multiple Cues in Learning

Manuel Miguel Ramos Alvarez Jaen, Andalucia, Spain

A Dual-Mode Model of Cooperation in Risk Management

Timothy Earle Bellingham, WA

Interattribute Correlations Influence Whether Decision Strategies are Option-based or Attribute-based

Barbara Fasolo, Gary H. McClelland and Katharine Lange Boulder, CO

The Role of Cue Intercorrelations in Students' Judgments of Course Interest

James A. Athanasou & Olu Aiyewalehimi Sydney, Australia

Why Are Some Judges Better Than Others?

Elise Weaver & Tom Stewart Albany, NY

Accuracy in Personality Judgments and a Commentary on Heuristics and Biases Questions

David Funder Riverside, CA

Brunswik Symposium Submitted for SPUDM

Mandeep K Dhami London, UK

Aging and Rule Learning

Gérard Chasseigne, & Peggy Lafon Tours, France

Linking Generalizability Theory and Social Judgment Theory

James Hogge Nashville, TN

Judgment Analysis, Think-Aloud Protocols, Cause Mapping, Image Theory and Neural Network Simulation

Ray Cooksey Armidale, Australia

Judgments about Union Representation, Crossing Picket Lines, Sexual Harassment Cases

Jim Holzworth Storrs, CT

Research in Patient Decision Support

Celia Wills East Lansing, MI

Physician Decision Making in Testing for Prostate Cancer

Paul Sorum Albany, NY

Scripts as an Alternative Framework for Describing Medical Decision Making

Robert Hamm Oklahoma City, OK

Physicians' Perceptions of Guidelines Regarding Prostate Cancer Testing

Junseop Shim Albany, NY

Culminations, 2000

Ken Hammond
Boulder, CO

The year 2000 has been an active one for me. My book Judgments Under Stress (Oxford) was published in January, on my 83rd birthday [almost]. Although I have not yet seen any reviews, I am pleased with this book because it fulfills a long-standing aim: to work out an application of Brunswikian theory to an important area of J/DM.

I chose the area of the effect of stress on judgment because it presents a central topic of both theoretical and practical importance that can hardly be over-estimated. In addition, it is one that I have long believed represented an excellent example of psychological research at its worst; that is, the repeated application of the "find the effect of a variable" paradigm, leaving every conclusion contingent upon a limited set of conditions.

So, in an attempt to demonstrate how Brunswikian psychology can produce a new, different, and productive approach, I present in this book a conception of stress derived from Brunswik's emphasis on the theory of constancy as "the essence of life" (1956, p. 23). From that standpoint, the disruption of constancy "presents a threat to the organism that induces not merely an affective response but a cognitive one" (Hammond, 2000, p. 69).

I insist that this point of departure is a solid one (in contrast to conventional work) based as it is on an empirical discovery (constancy) that represents not only a miraculous organismic achievement, but is apparent throughout the animal world.

But for those who would prefer to stand on the accomplishments of current and past stress research, I include for their evaluation an extensive annotated bibliography of this work, which I claim to be useless. In short, I offer a Brunswikian framework as a new point of departure for theory and research on stress-related judgments. Perhaps it will serve as a model for applications of Brunswikian theory to other areas, long smothered by conventional research methods.

A second publishing event occurred in 2000 that fulfilled another of my long-standing goals, no, dreams, and that was the completion of the preparation of Brunswik's English-language papers for publication. These are to be included in a book titled The Essential Brunswik: Beginnings, Explications, Applications (Oxford) (see http://www.brunswik.org/resources/ ebcontents.html).

The book manuscript is now in its final stages of production and will appear in 2001. It should provide ready access to original sources for Brunswikian researchers, something that has been lacking for nearly a half century. Those interested in Brunswikian theory and research will no longer have to be satisfied with secondary sources.

This work was carried out in collaboration with Tom Stewart, without whom this project would never have got off the ground, or seen the light of day. He did the hard, detailed work of gaining the cooperation of 29 authors, not including ourselves, and putting all of the 48 chapters together in a sensible form. Most important, he saw the preparation through to its conclusion, despite the obstacles and frustrations that efforts of this kind inevitably encounter. I suggest that at the next meeting of the Brunswik Society, Tom be given a medal in the form of a gold embossed lens model (of course).

As you probably know, all the royalties from this book will go to the treasury of the Brunswik Society.

I was happy to see that in 2000 Oxford saw fit to bring out a paperback copy of my Human Judgment and Social Policy book.

Currently, I am at work on a book manuscript to be titled Human Judgment in the Information Age: Getting Better - or Worse?" It is intended for the "trade" market. I am roughly half-way to reaching my goal.

Brunswik Society Website

Tom Stewart
Albany, NY

The revision of the Brunswik Society site is now complete. Please visit www.brunswik.org.

In addition, to redesign for easier navigation and future growth, several new pages have been added:

1. Brunswik's Woodworth: The story, the book, and images of selected pages (see "Resources").

2. Scenes from the 2000 meeting, in Berlin (see "Photographs").

3. A new essay by Bernhard Wolf (see "Notes and Essays").

4. Table of contents for the forth-coming volume, "The Essential Brunswik: Beginnings, Explications, Applications." (see "Recent Books by Brunswikians" under "Resources")

Oxford University Press is offering this book to Brunswik Society members at a substantial discount. A printable order form is provided. Royalties from the sale of this book will go to the Brunswik Society. Suggestions, contributions, and bug reports welcome.

Information Selection Dissertation and Brunswik Chapter

Michael Doherty
Bowling Green, OH

The following is an abstract of a dissertation currently being written up by John R. Leach, at Bowling Green State University:

"Information Selection in a Simulated Medical Diagnosis Task: The Effects of External Representations and Completely Natural Sampling"

Gigerenzer (1994) argued that cognitive biases could be made to disappear if tasks and problems were presented to subjects in formats analogous to natural sampling, rather than probability formats.

Several researchers (Gigerenzer and Hoffrage, 1995; Cosmides and Tooby, 1994; Betsch, Biel, Eddelbuttel, & Mock, 1998; Gigerenzer & Todd, 1999) have studied the accuracy of subjects on Bayesian inference problems presented in either a frequency format (i.e., analogous to natural sampling) or a probability format.

Results generally showed that frequency formats elicited higher proportions of solutions consistent with Bayesian solutions than did the probability formats. It was assumed that frequency formats correspond to natural sampling.

Gigerenzer and Hoffrage (1995) expressly stated that sequential acquisition of information by updating event frequencies without artificially fixing the marginal frequencies is what they refer to as natural sampling. However, the tasks they used did not have subjects sequentially update frequencies. Subjects were simply presented with summary information about events and outcomes.

In the present study, subjects sequentially sampled a task environment in a "completely natural" style. They were presented with a fictitious medical diagnosis scenario. This scenario offered two competing hypotheses about symptoms associated with two different fictitious diseases. The study was designed to determine if completely natural sampling is different than natural sampling as conceptualized by Gigerenzer (i.e., frequency formatted summary information), and if subsequent diagnostic and probability judgments differ. Diagnoses and judgments of probability were analyzed in relation to external representation and style of sampling (i.e., natural, completely natural, and controlled).

Subjects exposed to completely natural sampling were 100 percent accurate in diagnosing the disease. However, in a subsequent pseudodiagnosticity task, fewer than 50 percent were able to identify the information needed to calculate the Bayesian probability. Additionally, subjects in all conditions had trouble calculating probabilities associated with the diagnosis. Fewer than 30 percent reported the exact Bayesian probability. The results are consistent with the proposition that the cognitive processes involved in global judgments and diagnoses differ from those involved in analytical reasoning.

Ryan Tweney and I have just agreed to write a chapter on Brunswik for a book dealing with the history of thinking and reasoning. Given that I will be fully retired in two weeks, this is just the sort of scholarly project I should find of great interest in retirement.

In Defense of the Ecological Approach to Psychology

Kim Vicente
Toronto, Canada

My Brunswik-related research this year involved defending and explaining the value added by an ecological approach to psychology to a Newell & Simon-type audience. The abstract of the paper is as follows:

This article is part of an exchange concerning the contributions of the constraint attunement hypothesis (CAH) to the understanding of expertise effects in memory recall. My commentators claim that the CAH is not novel and that existing theories of the same phenomenon do not have the limitations that were attributed to them. In this reply, I argue that the CAH is the only theory of expertise effects in memory recall to adopt the abstraction hierarchy as a theory of the environment, a unique feature that has important theoretical implications.

Furthermore, other theories of this phenomenon focus on psychological mechanisms but cannot currently satisfy the burden of scientific proof required of process theories. Progress can be made by integrating the complementary advantages of existing theories into a unified theory that acknowledges the equally important roles of the organism and the environment.

The full reference is:

Vicente, K. J., "Revisiting the constraint attunement hypothesis: Reply to Ericsson, Patel, & Kintsch (2000) and Simon & Gobet (2000)," *Psychological Review*, 107, 601-608.

Integrating Social Judgment Theory and Simple Heuristics

Mandeep Dhami
London, United Kingdom

In January 2001, I will be taking up a post-doctoral position at the Department of Psychology, University of Maryland, to work with Tom Wallsten.

Over the last year, I have been working on a few ongoing Brunswik-related projects, and most are close to completion. One project that also fulfills a personal goal is the completion of my doctoral thesis entitled "Bailing and Jailing the Fast and Frugal Way: An Application of Social Judgement Theory and Simple Heuristics to English Magistrates' Remand Decisions." In this, I conducted three studies that address issues of interest to neo-Brunswikian researchers. These included decision maker-related questions concerning cue use, consistency, agreement, post-decisional confidence, and "insight." A main aim was to compare the relative validity of different models of the judgment process. The other questions concerned the effect of task characteristics such as information availability and time pressure on the decision maker. I also compared policies captured using what Brunswik (1956) would call a representative design, via an observational study, and policies captured using systematic design, namely hypothetical cases comprising an orthogonal cue set.

In England and Wales, magistrates (most of whom are lay judges) make decisions on around 98 percent of all criminal cases. Their decisions have huge ramifications for defendants, their families, the criminal justice system, and the general public. However, until now, they had escaped scrutiny from psychological research. Similar to past SJT research, I found that most magistrates demonstrated inconsistency in their decisions in a test-retest situation. The modal response was recorded for each of a set of cases, and each magistrate disagreed with this response on at least some cases. The cues that were found to influence magistrates' decisions according to their models differed from those magistrates directly reported using. Here, the use of extra-legal cues was under-reported. Despite these results, all magistrates reported feeling highly confident in their decisions.

Unlike most past SJT research, I compared alternative descriptions of the judgment process. Brunswik (1943, 1952) favoured the use of correlational statistics for describing the process of vicarious functioning. Following this, past SJT research has tended to use the multiple linear regression model to capture and represent individuals' judgment policies. These static, structural models depict peoples' judgments as being the product of a linear, compensatory integration of multiple cues that are differentially weighted. Also, the same cues are used in the same way on each case. As such, these models do not meet the criteria of psychological plausibility, flexibility and adaptability. Brunswik (1955, 1956) recognized that there may be alternatives to the correlational model, as did Hammond (1955, 1996).

Recently, simple, fast and frugal heuristics have been proposed as viable alternatives to the regression model (Gigerenzer & Goldstein,1996). These are precisely defined process models that comprise principles for information search, stop search, and decision making. Many are non-compensatory. I developed a model for binary categorization tasks, called the Matching Heuristic. In this model, few cues are searched in order of their validity, and search stops once a cue that points to a punitive decision is found. Only that cue is then used to make the decision. There is no integration of cues.

In each study, I compared the relative ability of three models (i.e., Matching Heuristic, a unit-weighted linear model and a differentially weighted linear model) to predict magistrates' decisions, both on the cases used to form the models and on a new set of cases. The results were consistent across the studies: the Matching Heuristic proved a better descriptor and predictor of magistrates' decisions, than the other models. This result held when magistrates were presented with an orthogonal cue set, and a representative cue set. It also held when magistrates worked as individuals, and when they worked as benches (small groups). Some of these results are already in press (Dhami & Ayton). In a study in the medical domain, the Matching Heuristic was found to do as equally well as a logistic regression model in describing English doctors' prescription decisions (Dhami & Harries, in press).

Although the simple heuristic proved descriptively and predictively valid, I recognised that it lacked prescriptive utility in the legal domain, where accuracy is not the main criteria by which to evaluate the quality of a decision. Packer's (1968) definition of the requirements of justice and due process, more closely resemble the workings of a regression model than a fast and frugal heuristic. I have been consulting with the British government on these findings and thinking of ways to improve performance. In their exposition of SJT, Hammond et al. (1975) stated that "social judgment theorists firmly believe that all students of human judgment should engage in research that will help provide better social policies and thereby increase our chances for a decent life on earth" (p. 306). I guess you could say I have been trying to do my bit.

Brunswik, the Berlin Way

Ulrich Hoffrage
Berlin, Germany

Almost all of my activities in 2000 that were related to Brunswik have been done together with my colleague Ralph Hertwig. Because he has already sent his report to the mailing list, I can keep my list short by just referring to his.

First, we organized this year's Society meeting - and let me once more use the opportunity to thank those of you who participated for making the effort to come and for your contributions.

Second, our Brunswikian report on the hindsight bias has been published this year in the JEP:LMC. The main idea is that a lack of direct memory on what we said or thought in the past can (and will) be compensated for by reconstruction based on the knowledge underlying our initial judgment.

Feedback on the correctness of this judgment leads to an update of the cue values we used for our original judgment. Note that in this process of updating, the cue-criterion relationship is reversed. The reconstructions from the updated knowledge base, in turn, lead to distorted "memories" - the so-called hindsight bias. (I wouldn't be too surprised if someone finds this short summary hard to understand, but I hope it gives you the feeling that this work is in the spirit of Brunswik and makes you curious enough to read the article.)

As Ralph already mentioned, we are now in the process of running simulations of this model - most of the work is done by Carola Fanselow, a graduate student in our group.

Third, Mandeep Dhami (and Ralph and I) have been - and still are - working on a review of the use of representative design in Social Judgment Theory research.

Fourth, we took the challenge of designing and testing a fast and frugal heuristic for a task at which regression is particularly strong: numerical estimation.

Last but not least, there is one thing that Ralph did not mention (at the time he sent his report, we didn't know yet). We are happy to announce that Science will publish in its Policy Forum Column an article by us (U. Hoffrage, S. Lindsey, R. Hertwig, G. Gigerenzer: "Communicating Statistical Information," scheduled for December 22nd, 2000). In this paper, we demonstrate that experts' and laypeople's difficulties in dealing with statistical information can be considerably reduced if the information is communicated in what we call "natural frequencies."

The link to Brunswik is that natural frequencies are the result of natural sampling, that is, of observing and counting cases as they occur in the environment, as one would do in a representative design. In the article, we describe what natural frequencies are in more detail, report two new studies in the domains of medicine and law, and discuss several applications and implications.

Max Planck, Columbia, Representative Design, and Two Fast & Frugal Heuristics

Ralph Hertwig
New York, NY

In 2000, my Brunswik-related activities included co-organizing (with Ulrich Hoffrage and Gerd Gigerenzer) the 2000 Brunswik Society meeting, which was held in Europe for the first time in the Society's history. (The meeting took place at the Max Planck Institute for Human Development in Berlin.) Although we three organizers are admittedly not impartial judges, we thought the meeting was a success thanks to many interesting contributions. In particular, we were happy to see attendees from many places throughout Europe who would normally not come to the meeting.

Second, Mandeep Dhami, Ulrich Hoffrage and myself are in the process of finishing a paper that reviews the use of representative design in Social Judgment Theory research. Mandeep, who is conducting the bulk of the work, presented some of the major results of this project at the 2000 Society meeting. Judging from some of the responses to her presentation, we expect the review to yield some challenging results, thus (we hope) encouraging a discussion of the future of representative design in the field.

Finally, in collaboration with colleagues at the Max Planck Institute, I have continued to work on some of the fast and frugal heuristics explored in the recent book authored by Gerd Gigerenzer, Peter Todd and the ABC research group (Simple Heuristics That Make Us Smart, 1999, Oxford University Press). In one project, we are further exploring a cue-based inference mechanism that we proposed as a model of a classic memory bias, namely, the hindsight bias (U. Hoffrage, R. Hertwig, & G. Gigerenzer, 2000, Hindsight bias: A by-product of knowledge updating? Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 566-581). Having already implemented this mechanism in a computer simulation, we are now attempting to find out to what extent it can predict previous findings in hindsight bias research and make novel, untested predictions.

We have also continued to work on an estimation heuristic called QuickEst, which exploits an ubiquitous environmental structure--J-shaped distributions--that characterizes a variety of naturally occurring phenomena, including many arising from accretionary growth. We are currently testing the performance of this heuristic in other environmental structures and are exploring the extent to which it can describe how people arrive at quantitative estimates.

P.S. This year I received a research grant from the German Science Foundation that allows me to spend two years abroad. In October, I moved to Columbia University where I am working in Elke Weber's lab. My new address is as follows: Ralph Hertwig, Columbia University, Schermerhorn Hall, 1190 Amsterdam Avenue, New York, NY 10027, Tel: (212) 854-4815

Interfaces and Time Pressure: Effects on Team Performance

Leonard Adelman, Sheryl Miller, and Cedric Yeo
Fairfax, VA

This is an abstract for the talk I gave recently at the Brunswik Society meeting in Berlin:

Relative Effectiveness of Different Interfaces to Ameliorate the Negative Effects of Time Pressure on Team Performance

An experiment was performed to investigate the relative effectiveness of a perceptually oriented interface, versus one providing cognitive feedback, to ameliorate the effect of increasing time pressure on the performance of hierarchical teams, which were represented conceptually by the multi-level lens model. The perceptually oriented interface was more effective than the one providing cognitive feedback because its visual cues helped teams maintain a high percentage of judgments under increasing time pressure. The cognitive feedback condition did not maintain high judgment accuracy, as had been predicted. Only the time pressure manipulation significantly affected judgment accuracy. A causal model using lens model equation parameters and Multi-Level Theory constructs (e.g., team informity and staff accuracy) showed that the time pressure effect was fully mediated by decreasing team informity. As team informity decreased with increasing time pressure, staff accuracy decreased (due to lower G) and leader accuracy decreased (due to lower G and Rs). These results suggest that time pressure effects on a team's judgment accuracy may be due more to a breakdown in information flow than a breakdown in judgment processing.

Base Rates from a Brunswikian Perspective

Phil Dunwoody
Athens, GA

I have just successfully defended my dissertation and am including the abstract below.

For my dissertation I, with the help of Adam Goodie and Robert Mahan, examined the phenomenon of base rate neglect from a Brunswikian perspective. That is, I examined base rate information usage as a function of its ecological validities.

In this study, we tried to borrow from both the heuristics and biases approach as well as the Brunswikian approach. The marriage proved successful, and we will be submitting this paper for publication shortly.

Below is the title and abstract.

"The use of base rate information as a function of experienced consistency and utility" by Philip Dunwoody, Adam Goodie, Robert Mahan.

The use of base rate information has been widely studied in decision making with the conclusion that people underweight or ignore base rate information when compared to a normative standard. This work extends the current body of research by demonstrating that base rate usage is moderated by the statistical characteristics of the base rate information. Two studies demonstrated that experienced base rate consistency and utility both affect base rate usage. Experiment 1 showed that participants use base rate information more when it is consistent than when it is inconsistent. Experiment 2 showed that when base rate consistency and utility are manipulated separately, participant decisions are mostly influenced by the utility of the base rates, and not the consistency. These studies demonstrate that base rate usage can be an adaptive response to environmental contingencies.

Situation Awareness

Alex Kirlik
Atlanta, GA

My student, Richard Strauss, completed a dissertation this year viewing the human factors problem of "situation awareness" (SA) from a Brunswikian perspective. Rich developed a framework in which SA is conceived as a set of relations between a performer's state of knowledge and an environmental situation. He conducted three experiments in the context of a naval submarine detection task in order to evaluate his approach. Stewart and Lusk's expanded lens model (ELM) was used for statistical analysis and modeling.

The first experiment evaluated whether perceptually augmented displays could be used to enhance SA and whether ELM parameters would be sensitive to this manipulation. The results of this experiment were mixed in terms of SA enhancement, but unequivocal when evaluated in terms of the ELM. Display augmentation improved participants' abilities to perceptually measure cue values but caused participants to display a significantly greater regression bias than participants using a baseline (unaugmented) display. The bias of the display augmentation group was toward overweighting situation specific cues at the expense of base rate information. This finding is in accord with the suspicion of some human factors researchers that the use of increasingly rich displays for decision support may cause performers to insufficiently attend to information from other sources.

In the second experiment, a fitted ELM model for each experimental participant was used to bias the presentation of display information in a participant-specific manner. This manipulation significantly increased environmental predictability for 13 out of 16 participants and increased the achievement of 8 of 16, leaving the other 8 unchanged. The biasing adjusted cue values to levels where the ELM model predicted the participant would render a perfect judgment, and had the effect of acting as a filter on environmental noise.

The final experiment focused on individual differences, in particular, the ELM's ability to diagnose the underlying differences between the highest and lowest performing experimental participants. In this task, high

and low performers did not differ in terms of regression or base rate biases, or in terms of task knowledge. However, high and low performers differed significantly in terms of both consistency of cue acquisition and consistency of information processing.

Strauss (2000). A methodology for measuring the judgmental components of situation awareness. Unpublished doctoral dissertation. School of Industrial & Systems Engineering, Georgia Institute of Technology.

Otitis, Multiple Cue Probability Learning, and Control Theory

Claudia Gonzalez Vallejo
Athens, OH

My Brunswikian research:

1) I have followed up on work on otitis media that started in Albany, NY, when I worked with Tom Stewart. In collaboration with him, Junseop Shim, Paul Sorum, Gérard Chasseigne, Maria-Teresa Sastre, and Etienne Mullet, we recently rounded up a study that looked at diagnosis and treatment decisions of physicians in the U.S. and France, plus the responses of parents. Main results showed that both sets of physicians relied heavily on examination cues to make their judgments. Physicians within groups disagreed, but in comparison to each other, both French and U.S. doctors appeared to use similar strategies. There was also a lot of similarity between parents and physicians in terms of diagnosis and treatment judgments, even though the groups held different attitudes. In addition, some further analyses of these data that I performed show that the treatment decisions are not solely based on the information used to make the diagnoses, but that additional variables influenced treatments.

2) The other project is one that is still in the conception stage and deals with using control theory (Powers, 1978) to create a dynamic model of the person and the task in MCPL. This work is in collaboration with my colleague at Ohio University, Jeff Vancouver, who has recently programmed a tentative simulation model, which relies heavily on the notion that the human cognitive system, like the physiological system, tries to maintain equilibrium via negative feedback loops. My next task is to get data sets, and here is where you all can help me and guide me where to find the data, in order to test the dynamic predictions that the current version of the model(s) make (we can rather talk of a family of models)

Optimal Hierarchical Command Styles in Dynamic Decision Tasks

Alexander Wearing
Melbourne, Australia

This is work being carried out by Julia Clancy, Glenn Elliot, Tobias Ley, Jim McLennan, Mary Omodei, Peter Taranto, Einar Thorsteinsson and Alexander Wearing.

Tasks involving dynamic decision making, such as fire fighting and medical emergencies, are commonly distributed among a number of people.

The organizational structure is typically hierarchical in nature, with tasks and responsibilities divided in a structured way among incident commanders and subordinates. However, the optimal way to divide the responsibility of decision making among team members is not obvious.

Should commanders make all decisions and communicate actions for the subordinate to carry out? Or is it better for decision making responsibility to be shared, with commanders communicating their intentions and subordinates then deciding on appropriate actions and carrying these out? This is fundamentally an issue of the relative effectiveness of different command styles, which create different distributions of task responsibilities. We have addressed the issue by using computer-simulated fire fighting tasks, typically undertaken with teams of three, one commander and two subordinates. The results indicate that the teams where commanders communicate their intent perform significantly better than the teams whose commanders communicate specific actions. Depending on the experimental condition (action or intent) commanders differ with regard to activities such as predicting the development of fires, monitoring wind direction (current and forecast), fire front prioritization, allocation of appliances, and moving appliances. There is little evidence of consistent individual differences.

Uncertainty, Confidence and the Lens Model

Ann Bisantz
Buffalo, NY

This year several graduate students and I have been engaged in a variety of projects related to decision making.

Richard Finger completed an M.S. thesis in which he investigated the use of degraded and blended icons to convey uncertainty regarding an object's identity. The thesis consisted of two studies in which icons were used to convey the probability that the identity of a radar contact was hostile or friendly.

A pilot study first investigated whether participants could sort, order, and rank icons from seven sets intended to represent different levels of uncertainty. For example, one set used an "X" to indicate a 100 percent chance of the entity being hostile, and an "0" to indicate the opposite. As the chance approached 50 percent, the icons used were more and more distorted (made more "fuzzy" using a pixelating function in a graphics package).

Results from the pilot study indicated that sets of degraded or blended icons intended to represent levels of situational uncertainty could be ordered and rated in a manner similar to expectations. Additional results from the pilot study indicated a framing effect on performance: Participants' interpretations of displayed information became less ideal in a negatively framed context (that is, when those icons were described as representing hostile rather than friendly entities). Finally, results from the pilot study were similar across icon sets, indicating that experimental results were not necessarily specific to a particular icon form.

Three icon sets were selected for further study in a decision making experiment, in which participants had to identify objects as either hostile or friendly. Participants saw a simulated radar screen in which unidentified contacts and probabilistic estimates of their identities were depicted in one of three ways: with degraded icons and probabilities, with non-degraded icons and probabilities, and with degraded icons only.

Results showed that participants using displays with only degraded icons performed better on some performance measures, and as well on other measures, than conditions where degraded icons were annotated

with numeric probabilities, where information regarding uncertainty was conveyed only via numeric probabilities, or where numeric probabilities were mapped to the icons in the task instructions.

These results are significant because they indicate both that people are able to understand uncertainty conveyed through such a manner and, thus, that the use of distorted or degraded images may be a viable alternative to convey situational uncertainty.

This work will continue through both an NSF-funded grant, and an association with the Center for Multi-source Information Fusion at the University of Buffalo.

Chang-soo Nam also completed an M.S. thesis in which he studied the relationship between task characteristics and confidence in judgments. Nam used a pavement judgment task, in which participants had to make a decision regarding a type of repair given characteristics and a photograph of a pavement crack. Participants were provided with cognitive feedback regarding their performance across four sessions.

A common effect in studies of confidence is that participants tend to be more overconfident in their judgments (that is, more confident than their performance indicates) on tasks that are more difficult. This result was confirmed in the present study - but in this case, assessments of task difficulty were made a priori, and objectively, by creating experimental conditions across which judgment predictability (Re) varied. That is, tasks were grouped into two conditions with high and low Re. Additional analyses indicated that participants who were more overconfident had lower values of G (linear knowledge) indicating poorer adaptability to the linear structure of the environment than those who were not overconfident, but had similar levels of consistency (Rs).

One interpretation of these results is that participants who are overconfident may be reflecting, in their self-assessments of confidence, their belief that they are making judgments consistently.

Additionally, two other students are pursuing lens-model related work. Gordon Gattie is beginning data collection for a study applying aspects of cognitive feedback to train dental students in classifying oral cancers, and Pratik Jha is applying the multi-variate lens model to describe fault diagnoses in process control.

Self-regulation, Feedback and Leadership Skills

Shawn Noble
Boston, MA

Below you will find an abstract of my dissertation, which I successfully defended this past October:

Teaching managers how to think is one of the most important determinants of developing leaders who are versatile, adaptive, and flexible. This notion of "how to think" is defined by the Army as a conceptual skill.

Previous research by Noble and Fallesen (2000) has shown that conceptual skills can be divided into three categories: simulation, synthesis and self-regulation. Specific focus was taken in two studies to gain a better understanding of self-regulation thinking skills.

In Study 1, 150 undergraduate students were studied to explore a model that considered the relationship between variables that are thought to be related to a self-regulatory process. The variables included locus of control, self-efficacy, and goal-setting habits. In addition, the model considered how well self-efficacy for

improving leadership and goal-setting habits predict academic achievement goals and behavioral intent. Results from Study 1 showed mixed support for the model.

Study 2 used 277 military officers to expand the model from Study 1 to explore the relationship between self-insight and self-regulation. The Study 2 model was designed to address: 1) How feedback influences leadership goals and attitudes of leadership assessment; 2) The relationship between locus of control, perceived feedback consistency, self-efficacy, goal-setting habits, system reactions, and perceived constraints; 3) The relationship between behavioral intent and self-efficacy, goal-setting habits, system reactions, and perceived constraints.

Support was not found for the original model; however, based on the findings in Study 1 and Study 2, an alternative descriptive model was proposed that draws upon a different association between each of the variables and supports the idea that individual difference variables play an important role in the self-regulatory process.

The research provided insights on the AZIMUTH multi-rater feedback process utilized by the U.S. Army to develop officer leadership. Results showed that locus of control and self-peer rating agreement play a key role in determining how feedback is interpreted. In addition, the research conducted here led to the development of several scales that can be used for future exploration of locus of control and self-efficacy.

The Use of Configural Information in Multi-Cue Judgments by Individuals and Groups

Scott Tindale
Chicago, IL

Here is an abstract for a talk I am giving at the Judgment and Decision Making Society meeting this year in New Orleans - a similar paper will be a poster at the Society for Personality and Social Psychology meeting in February. My co-authors are Elisabeth Anderson, Amanda Dykema-Engblade, Helen Meisenhelder, Catherine Munier, and Andrea Krebel.

Research has shown that groups tend to out-perform individuals in terms of accuracy on multi-cue judgment tasks. Research has also shown that configural information (information imbedded in the pattern of the cues, i.e., interaction effects) is more difficult to learn from feedback than is information associated with the simple cue-criterion relations. However, over trials, individuals do learn to use configural information.

This study compared individuals and groups in their ability to learn to use configural information. Individuals and groups rated 60 job candidates based on two pieces of information. They received accuracy feedback after each judgment. Two types of configural information were addressed: disjunctive (best score is more important) and conjunctive (worst score is more important). Two job types were used: math tutor and computer analyst. The job and cue descriptions were designed to produce expectations of either a disjunctive or conjunctive relation between the cues and criterion.

Results indicated that groups outperformed individuals but that they were no better than individuals at using the configural information. Conjunctive information was easier to learn than disjunctive information, particularly when strong conjunctive expectations were in place. The superiority of groups appears to stem from their more consistent use of the linear cue-criterion relations as compared to individuals.

Dealing with Conflicting Information

Kathleen L. Mosier, Jeffrey Keyes, and Roberta Bernhard
San Francisco, CA

Omission and commission errors resulting from automation bias, the tendency to rely on automated cues as a heuristic replacement for vigilant information seeking and processing, have been documented in professional pilots and students, in one- and two-person crews. Underlying causes of omission errors have been traced in part to vigilance issues, as crews who are monitoring flight progress and system status often "miss" events that are not pointed out to them by automated systems. Causes of commission errors are harder to track. It has been hypothesized that they may be related to a desire of pilots to "take action," as proactivity has typically been associated with superior crew performance. In this work, which focused on regional, or Part 135 operations, the decision involved in choosing among sources of information was investigated via an ASRS (Aviation Safety Reporting System) analysis, and also through a paper-and-pencil scenario study.

Data for the ASRS study were obtained from the ASRS CD containing the database for the years 1994-1998. Using several broad search criteria words, such as "automation" (or specific automated displays or instruments) and "conflict," we created a preliminary sample of 1,200 reports that were potentially relevant to our study. Each of these was screened for appropriateness, and we identified 189 ASRS reports in which automation was involved. Incidents were coded with respect to the sources of information that were cited concerning the critical incident, whether the sources provided consistent information or were in conflict, and how the incident was resolved. We were particularly interested in incidents involving conflicting information from different sources - and found that most of these incidents (N=24) were traffic incidents, and involved a conflict between TCAS information and some other source (ATC or visual cues). Traffic incidents were also most often cited as involving high risk. Analyses indicated that, when TCAS information entailed taking evasive action, crews typically followed TCAS recommendations - even when visual information contradicted the need for the maneuver. These incidents supported the notion of a "take action" tendency.

For the scenario study, 125 regional air pilots were asked to respond to a packet of scenarios. Each scenario conveyed a situation involving conflicting information from two sources - an automated source + either a human source or a traditional indicator. Information from one source suggested making some change (action); information from the other source suggested maintaining status quo. Seven of the scenarios were matched between packets - that is, the same scenario was manipulated so that, in one packet, the information from the automated source suggested action, and in the other packet, the information from the other source suggested the same action. One scenario contained conflicting action recommendations - an automated source suggested that one of two engines was on fire; traditional indications suggested that it was actually the other engine that was damaged. Pilots saw only one version of each scenario. They were asked to choose between two decision options, and assess their confidence level, as well as the risk involved in the scenario.

We found no systematic evidence of a preference for automated information - in fact, in none of the scenario pairs was automated information followed across packets. Rather, we saw a pronounced scenario effect; that is, in most scenarios there was high agreement across packets on the preferred option, the risk level of the scenario, and the confidence with which pilots chose an option. For the pair of scenarios that contained conflicting engine fire indications (which engine was on fire), pilots most often believed traditional indicators. We did not find evidence of a systematic preference for action (which was, in most cases, the more conservative option), although the higher the estimated risk of a scenario, the more likely pilots were to choose action, and the more confident they were in their choice.

Results of this study are encouraging in that they suggest that we may be able to impact automation bias if we train pilots early enough in their careers to evaluate automated cues in context with other cues. However, we need to be cautious about generalizing from the paper-and-pencil venue. This format provides information differently than it is shown within the cockpit, and allows the information to be processed in a less biased and more analytical way. Additionally, we have previous evidence that, when encountering a situation in an actual or simulated aircraft environment, pilots do not always do what they say they would do. Follow-up studies will be required to determine if results of the paper-and-pencil study will hold in other venues.

Integration of Multiple Cues in Learning

Manuel Miguel Ramos Alvarez
Jaen, Andalucia, Spain

My research deals with predictive processing in binary situations when these situations present more than one potential predictor of consequences.

The main result obtained in our laboratory up to now is that predictors are processed independently in some situations such as the blocking one. These results do not agree with the competitive mechanism as shown by related literature. Psychologically speaking, the discounting or competition principle cannot explain our results.

In our previous research we proposed an Information Integration Model based on Norman Anderson's theory and also related it to the Brunswik Lens Model. Our theoretical proposal takes its roots from two parallel regression equations, as in the lens model: One for the ecological structure, and the other for the subjective system. Such equations allow the estimation of multiple regression parameters for each of the predictors included in the situation. These estimations are carried out through the computation of the focal cue validity (or pairwise correlation) relative to the validities of the other potential cues.

The linking function between the ecological-environmental structure and the cognitive-subjective systems incorporates the subject's beliefs or assumptions about the causal texture. For instance, in the blocking context the person may have the belief that the cues are independent of each other, and this fact would lead to a simplified subjective estimation only based on the crude validity (the cue-criterion or cue-consequence correlation).

In addition, relevant information to estimate the regression index of each predictor could be integrated according to different rules. We propose a formal model with two types of Information Integration Strategies, including every possibility up to date. Either the person uses a regression rule based on absolute frequencies (sub-optimal heuristic strategy) in which the weight and the information sign may differ, or the person uses a rule in which the information is processed in a relative or probabilistic way. At the same time, each rule can be subdivided according to the person's assumptions. For instance, if one kind of causal link is assumed, the model is slightly different than if the assumption is the opposite one.

Within this framework, we have carried out new research related to the processing strategy and research related to complex stimulus situations:

A) In the former research area, we have manipulated the contingency, the expected associations between cues and outcomes (high positive expectancy or Null expectancy), and the type of information (Symmetric or Asymmetric). We have shown the usefulness of all the regression strategies in our model.

We found that a high percentage of the people used heuristic rules in all the experimental conditions. With respect to the Normative Strategies, a high percentage of the subjects assumed a direction of regression opposite to the conventional, X-to-Y. There was also evidence of a very low percentage of strategies that did not take the appropriate information sign, according to the association established beforehand in the predictive situation.

In addition, our research work helped us establish the subjective weight pattern of the different information types.

Only some of the experimental conditions were adjusted to a uniform weight distribution. Some other conditions followed a non-uniform pattern proposed in the investigation of causal-predictive learning. Still some other conditions deviated from the two patterns above.

B) Regarding our research about complex stimuli, we have extensively explored all those conditions that have an influence on the competition principle.

The competition is only present when the situation allows the direct comparison of diverse types of experiences, i.e., a relative validity paradigm, and when the context is a causal one. In the blocking paradigm, contingency judgments do not follow the competition principle unless the experiment induces the belief that the cues are associated with each other.

A Dual-Mode Model of Cooperation in Risk Management

Timothy Earle
Bellingham, WA

According to the Japanese social psychologist, Yamagishi, failure to clearly distinguish trust from confidence is a "major source of confusion surrounding discussions of trust." During the past year, we have tried to clarify the distinction between trust and confidence, first, by reviewing a large number of studies drawn from many disciplines and, second, by constructing a general, dual-mode model of cooperation based on trust and confidence. In this brief report, I outline the elements of our model. An expanded treatment, with references, is available.

We assume, as many dual-mode theorists do, that the two processes operate simultaneously and that interaction between the two consists primarily of associative thinking affecting rule-based thinking, though effects in the other direction are also possible; i.e., the associative system can be overruled. One immediate implication of this is that trust should affect confidence, but only when confidence is rule-based ("rational," "analytical," etc.). But when confidence is assumed (i.e., associative), there is no social uncertainty, and trust is irrelevant. Thus, whenever trust is in play, it should affect judgments of confidence.

Our model depicts two pathways to cooperation, the upper via trust, the lower via confidence. At the far left of the model, the information perceived by a person is divided into two types, that which is judged to be relevant to "morality" and that which is judged relevant to "performance." This division of information, although central in studies of impression formation, has been overlooked in most studies of trust and confidence. The importance of this distinction is demonstrated, first, by studies that show that persons tend to organize impressions of others along two dimensions, social desirability (morality) and intellectual desirability (performance), and, second, that morality information tends to dominate performance information. By "dominate" we mean that, to an observer, morality information is more important and that it conditions the interpretation of performance information. For

example, given positive morality information, negative performance is judged much less harshly than it would be if the morality information were negative. The elements of our model are aligned in parallel pairs for trust and confidence:

a. Perceived "Amplitude" of (Morality/Performance) Information. The judged degree to which the given information has (morality/performance) implications.

b. Perceived "Valence" of (Morality/Performance) Information. The judged degree of positivity/negativity of the given information. (a. and b. combine to form c.)

c. Attributed (Values/Performance). The (values/performance) attributed by the observer to the other.

d. Salient Values/Salient Performance History. In the case of values, these are the values that are currently salient to the observer--which may be the product of existing social trust relations. In the case of performance, this is whatever history of relevant performance that is currently available to the observer.

(c. and d. combine to form e.) e. Value Similarity/Perceived Performance. Value Similarity is the judged similarity between the observer's currently salient values and the values attributed to the other. Perceived Performance is the observer's interpretation of the other's performance; note that this is a product not only of c. and d., but also of Social Trust, element g., below.

f. General Trust/General Confidence. General Trust is generalized interpersonal trust, the belief that most people can be trusted. General Confidence is the performance-based counterpart of the values-based General Trust: the belief that things in general are under control, uncertainty is low, and events will occur as expected.

(e. and f. combine to form g.) g. Social Trust/Confidence.

h. Cooperation. Any form of cooperative behavior between a person and another person or group of persons, or between a person and an organization/institution.

Among the key features of our dual-mode model of cooperation are these: 1) It shows how social trust is based on morality-relevant information, while confidence is based on performance-relevant information; 2) It shows how, in times of low social uncertainty, when morality information isn't relevant, social trust doesn't play a role in cooperation; 3) It shows how social trust becomes important in times of uncertainty, when morality information is relevant; and 4) It shows how social trust affects judgments of confidence via effects on perceived performance.

A number of important, testable hypotheses can be derived from this model; for example, a reinterpretation of "trust asymmetry," a widely-accepted "fact" of trust--that trust is hard to win, but easy to lose. In terms of our dual-mode model of cooperation, "trust asymmetry" collapses the two modes to one: Instead of morality information and performance information, there is simply information; instead of trust and confidence, with trust affecting confidence, there is simply trust. Most seriously, trust asymmetry ignores the role of existing attitudes and beliefs, which enter our dual mode model as Salient Values and Salient Performance History.

Interattribute Correlations Influence Whether Decision Strategies are Option-based or Attribute-based

Barbara Fasolo, Gary H. McClelland and Katharine Lange
Boulder, CO

This year our research has focused on how people make decisions amongst different options in the presence of "unfriendly" environments. In unfriendly environments, choice attributes are negatively related, and decision makers need to make trade-offs to form an overall evaluation of each option. But, trade-offs are hard to make. Our goal was to understand if decision makers ignore these trade-offs and use the same simple strategies that they would adopt in the presence of friendlier positive inter-attribute correlations, or if they adjust to the "unfriendliness" of the decision environment by using different strategies.

In two experiments, we presented decision makers with choices on Web-based Information Display Boards that were either friendly (characterized by positive inter-attribute correlations) or unfriendly (characterized by negative inter-attribute correlations). The change in correlation was manipulated between subjects in the first experiment, and within subjects in the second experiment. In both experiments, the change in correlation was implemented by simply changing the attribute values, while keeping the other choice variables constant.

We found that decision makers perceive attribute correlations in real choice situations and adjust decision strategies accordingly. In particular, when inter-attribute correlations are positive and the choice environment is friendly, decision strategies are simpler and attribute-based. The decision is judged easier, the attribute values are considered more predictable, the information search is not thorough, but is systematic. When correlations are, instead, negative and the choice environment is unfriendly, decision strategies are more effortful and option-based. The decision is judged difficult, the attribute values are more unpredictable, hence the information search is more erratic and more thorough.

This adaptivity to changes in correlation seemed to happen instantly upon "feeling" the friendliness of the choice problem. Participants who first encountered unfriendly choices were right away more option-focused than participants who first encountered friendly choices. This suggests that decision strategies are affected "online" by the correlation structure of the attribute set at hand, and not as much by pre-existing beliefs about how these attributes are usually interrelated in the real world (e.g., larger size goes with more weight).

Overall, decision makers were "fast" (decisions among five digital cameras described on eight attributes were made in less than 30s, too little time for making complex computations), "frugal" (not all information was explored, especially in friendly environments), but were not unilaterally attracted to shortcuts based on attributes. In unfriendly environments decision makers rather converged toward more effortful and option-based strategies.

In sum, we believe these two experiments provide evidence that attribute correlation powerfully affects information search strategy and ultimate choice in ways that are not always "non-normative."

This research reinforces an optimistic view of decision makers able to flexibly adapt search and decision strategies according to the structure of their decision environment.

Our next step is to see whether the same results hold in a real decision environment - the Internet - where an increasing number of users naturally encounter "friendly" and "unfriendly" information display boards and make consequential decisions from them.

The Role of Cue Intercorrelations in Students' Judgments of Course Interest

James A. Athanasou & Olu Aiyewalehimi
Sydney, Australia

A person's interest in a subject acts as a key factor in many models of learning. Typically educational achievement and satisfaction are inferred from ratings of student interest, but these ratings involve perceptions and subtle decision-making about oneself. People employ such judgmental processes regularly in education and, in the case of interest, they may be based on the subject, the learning situation or many other factors. The purpose of this program of research is to study how people make judgments about how interested they are in a subject.

In the first phase, a case study of student judgment was undertaken that involved two undergraduate students from the University of Technology Sydney. Their judgments of interests were based on information that described six aspects of a subject rated from very low (0) to very high (9). The six aspects (cues) were whether the subject was challenging, fascinating, the quality of teaching, the usefulness of the text, the quality of the facilities, and the extent to which theory and practice were related. Participants were presented with 60 separate profiles comprising random values of the six cues. They studied the six cues in each profile and then judged their own personal level of interest on a scale from 0 (no interest) to 9 (very high interest) based on the information contained in each profile. Profiles (N=15) were repeated to determine consistency of judgment of interests. Results were analyzed using a single lens model that describes human judgment of interest. Results showed individual differences and complexity in judgment with an overall emphasis on the quality of teaching.

The second phase of the research program has investigated the role of cue intercorrelations in the judgment of interest. Students were presented with 75 profiles containing the same six cues, but in this case the intercorrelation between cues was varied. Four separate groups of adult students (N=28) were tested using cues that correlated 0.0, 0.3, 0.6 and 0.9. Students were presented with 75 profiles including 15 randomly selected and repeated profiles to determine test-retest reliability of judgments.

While previous research has investigated the effects of multicollinearity in multiple cue probability learning within a double lens model, this has been in contexts involving fewer than six cues and never in single lens models. However, much human judgment occurs under conditions that are consistent with the single lens model, that is, where the criterion is difficult to measure or unavailable.

Preliminary results indicate marked differences between groups in the multiple correlation of cues with the judgment and the cue utilities. The effect of increasing cue intercorrelations was to increase multiple correlations and cue utilities. Multicollinearity of cues in a single lens model affected judgments in a predictable and monotonically uniform manner. A summary of the results is available upon request.

Further studies that are planned for Phases 2, 3 and 4 of the project include focus on varying the content of the cues (cue labels).

Why Are Some Judges Better Than Others?

Elise Weaver & Tom Stewart
Albany, NY

We presented a poster at the Society for Judgment and Decision Making meeting in New Orleans describing pilot work (37 subjects) for our NSF-sponsored project on individual differences in judgmental skill.

We hypothesized that predictors of judgment accuracy in a multiple cue probability judgment task would include skill in cue probability learning as measured by Chasseigne et al.'s MCPL task. In addition we tested whether the following would also predict accurate judgment: crystallized intelligence, fluid intelligence, and coherent judgment (as measured by the Linda task, avoidance of violation of probability rules in subjective probability estimates, and the Wason selection task).

While our initial model was not a good fit, the following paths were supported: Fluid intelligence predicted judgment accuracy, and MCPL skill also predicted accurate judgment, over and above the contribution of fluid intelligence.

In contrast, scores on tests of coherence did not predict judgment accuracy. In addition, we found it necessary to separate different kinds of coherence tests because these tasks were associated with different kinds of intelligence (Wason with crystallized intelligence, and probability rules with fluid). Finally, we found that a frequency format of a Bayes task was correlated with accuracy, though not tested in our structural equation model.

We are somewhat encouraged in our hypothesis that judgmental accuracy is not wholly a function of intelligence, but is also related to skill at MCPL. We were surprised, however, that coherence tests were unrelated to judgment accuracy, since we asked people to make probability judgments. These are only preliminary results, and we need to refine our measures and increase the sample size.

Accuracy in Personality Judgments and a Commentary on Heuristics and Biases Questions

David Funder
Riverside, CA

Our lab, which we now call the Riverside Accuracy Project, continues its research on the process and moderators of accuracy in personality judgment.

To this end, we are in the midst of a new round of major data gathering that includes three-person group interactions recorded on videotape, self-descriptions of personality, peers' descriptions, a clinical interview, and a vast number of personality inventories including the MMPI.

We operationalize accuracy in personality judgment as the convergence among these data sources. For more information, please consult our web page, which is: www.psych.ucr.edu/faculty/funder/rap/Rap.htm

On another Brunswikian front, I recently prepared a commentary on an article soon to appear in Brain and Behavioral Sciences. The article presents evidence of individual differences in performance on Kahneman-Tversky type brainteasers of the sort typically used to argue for the presence of "systematic irrationality." The fascinating fact -- of which the authors of the article make too little, in my opinion -- is that performance on different heuristics/biases tasks are intercorrelated, and correlated with SAT scores! In fact, a psychometric view of these tasks shows that they correlate with each other and with the SAT total score about as well as an actual typical SAT item.

In my commentary, I observe that the presence of "difficult" SAT items -- that is, items that most test-takers get wrong -- has not, as far as I know, ever been used to argue that people are systematically irrational. And yet the existence of the heuristics/biases items, which most people also get wrong but which, like hard SAT items, some people consistently get right, has been used to argue exactly that point. The existence of stable individual differences in performance on heuristics and biases problems means that they amount to little more than demonstrations that some kinds of problems are hard to solve, which SAT has known for years. This still might be useful information, but in this light any claim evaporates that these problems demonstrate fundamental flaws in human cognition. Along with this claim evaporates the very reason the heuristics-and-biases approach became so famous in the first place.

I would be happy to e-mail anyone interested in a copy of this commentary.

Brunswik Symposium Submitted for SPUDM

Mandeep K Dhami
London, UK

At the last July meeting of the Brunswik Society in Berlin, we planned to make a greater effort to raise the awareness and profile of Brunswik-related research amongst the general Judgment and Decision Making community.

As planned, I am organizing a symposium to be submitted for possible participation at the 18th SPUDM (Subjective Probability, Utility and Decision Making) Conference which is organized by the European Association for Decision Making. The conference takes place once every two years. In 2001, it will take place between 20th - 22nd August at the University of Amsterdam, The Netherlands. For more information about the SPUDM conference, see - http://www.eadm.org/eadm-spudm.html

The theme of the symposium will be Ecological Rationality and the Effect of Task Conditions on J/DM. Contact me if you need any further information.

Mandeep K Dhami, Department of Psychology, City University London, EC1V 0HB, UK. tel: +44 020 7477 0155 fax: +44 020 7477 8581

Aging and Rule Learning

Gérard Chasseigne, & Peggy Lafon
Tours, France

We have been pursuing the research program initiated in 1997, aimed at examining the effect of age on learning in complex situations. The present work concerns aging and rule learning. It is aimed at examining the effect of age on the ability to learn multiplicative combination rules. In such rules, cue levels do more than simply add their effects to determine the criterion values; the combination rule is a multiplicative one. To our knowledge, there has been no work on such a topic.

The specific situation studied was the learning of the multiplicative relationship between daily tobacco intake, daily alcohol intake and risk of cancer. This situation was chosen as a function of three criteria. Firstly, we wanted to use a concrete situation, that is a situation elderly participants could consider as relevant to everyday life: tobacco smoking, alcohol drinking, and cancer suffering are life events. Secondly, we wanted to use a situation for which a multiplicative combination model existed: A multiplicative type model of the relationship between cigarette dose, alcohol dose and risk of cancer of the esophagus has been proposed (Tuyns, Péquignot & Jensen, 1977). Thirdly, we wanted to use a situation for which it had already been demonstrated that the combination rule implemented by naive participants was not multiplicative: It has been repeatedly demonstrated that naive participants eschew to apply a multiplicative rule when asked to estimate the risk of cancer associated with given alcohol-tobacco consumption combinations. They uniformly apply a disjunctive rule (Hermand, Mullet & Coutelle, 1995; Hermand, Mullet & Lavieville, 1997; Hermand, Mullet, Sorum & Tillard, 2000). That is, they consider that indulging in only one of these two behaviors represents a maximum health risk.

Our overall hypothesis was that elderly people, more than young people or mid-adults would experience difficulties in learning to apply a multiplicative rule. This hypothesis was based on the proposition made by Chasseigne, Mullet and Stewart (1997), and Chasseigne, Grau, Mullet and Cama (1999) that the differences between young and elderly people in probabilistic function learning are mainly related to flexibility of functioning (changing from one default combination hypothesis, here the disjunctive rule, to another combination hypothesis, here the multiplicative rule). Specifically, we expect that (a) young people would pass gradually, through learning, from the use of a disjunctive rule to the use of a multiplicative rule, and (b) elderly people would keep using the disjunctive rule, despite feed-back.

A total of 86 individuals (33 males and 53 females) participated in this experiment (23 young adults aged 18-25, 22 employed persons aged 40-50, 20 retired persons aged 65-74, and 21 older persons aged 75-90). None was institutionalized.

The materials consisted of one set of 25 cards (21 x 3.5 cm), each showing two cue values in the form of tobacco consumption levels (0 pack, 1/2 pack, 1 pack, 1 and 1/2 packs, and 2 packs a day), and alcohol consumption levels (0 glasses, 2 glasses, 1 bottle, 1 and 1/2 bottles, and 2 bottles a day). The criterion value (1-100), which expressed the risk level associated with the tobacco and alcohol consumption levels, was written at the back of each card. This value was computed from Tuyns, Péquinot and Jensen's (1977) rule.

The subjects were told that their task was to forecast the risk of esophageal cancer associated with the various consumption levels. They were provided with the 25 different vignettes. The actual value of the criterion was displayed on the back of the same card as outcome feedback (OFB). Subjects were asked to learn the relationships between the levels of the two indicators and the overall risk level. The subjects were shown six blocks of 25 trials. The first block was a familiarization block without OFB. The second, fourth, and last blocks

served as policy capturing blocks (no OFB). The third and fifth blocks were learning blocks in which OFB was provided to the subjects. The experiment was self-paced, and subjects completed the task individually. They took about 1-1.5 hours to complete the experiment.

Our hypotheses were well supported by the data. Before receiving feedback, the participants implemented a disjunctive rule. After receiving a limited amount of feedback (first learning block), these participants had already learned to use a multiplicative rule. Even after receiving a massive amount of feedback (two learning blocks), however, the elderly people still showed difficulties using the multiplicative rule. These results strengthen the proposition put forward by Chasseigne, Mullet and Stewart (1997) and Chasseigne, Grau, Mullet, and Cama (1999) that the differences between young and elderly people in function learning are mainly related to flexibility of functioning (changing one hypothesis to another).

Linking Generalizability Theory and Social Judgment Theory

James Hogge
Nashville, TN

This year I have been putting the final touches on a manuscript (with Steve Schilling) entitled "Multilevel Judgment and Reliability Analysis: Hierarchical Linear Models as a Bridge Between Generalizability Theory and the Lens Model Equation." The following is an overview of our paper:

Although generalizability (G) theory and the lens model equation of social judgment theory (SJT) are two distinct analytic models of judgment that have developed along seemingly separate paths, they share common theoretical underpinnings provided by Egon Brunswik's probabilistic functionalism.

In this paper, hierarchical linear models are shown to be a promising analytical tool for combining these two previously divergent approaches into a single coherent framework - a framework simultaneously capable of assessing the dependability (reliability) of judgments and modeling diversity in judgment models across judges. In effect, G-theory is extended to take into account the information (cues) upon which judgments are based, and the lens model equation's pairwise consideration of interjudge agreement is supplemented by an analytic framework that considers all judges simultaneously, both individually and collectively.

Scientific paradigms begin as theoretical and philosophical orientations directed toward solving particular questions of interest to their proponents. Methodological developments, including statistical methods, arise as a means for implementing a particular research paradigm. But as an approach matures, it can become limited by its methodology in the sense that the methodology becomes inadequate to address important issues. We argue that this is the case with both SJT and G-theory.

Aggregation of ideographic results and assessing the adequacy of linear models of judgment are two important issues for SJT; however, up to now, the standard SJT technique of ideographic multiple regression has provided no coherent methodology for addressing these issues. Similarly, the standard G-theory technique of random effects ANOVA has been inadequate for addressing the issue of facets nested within persons.

But just as paradigms can be limited by their methods, outside methodological developments can serve to expand individual paradigms and unite seemingly disparate paradigms. Recent advances in statistical methods have lead many researchers to consider the utility of implementing a hierarchical approach to the analysis of linear models in a variety of contexts in psychology, social policy, and education. By applying hierarchical

linear models in the context of human judgment, we see potential for the expansion and unification of two heretofore disconnected approaches for the analysis of human judgment data: SJT and G-theory.

Specifically, we demonstrate how HLM analyses supplement separate analyses within the frameworks of G-theory and SJT in several important ways:

1. Judges are simultaneously described at the ideographic and nomothetic levels; thus, the tension between the classic ideographic Brunswikian emphasis and the nomothetic orientation of G-theory is resolved.

2. The HLM analysis easily deals with both linear and quadratic relationships between the cues and judges' ratings. There is no need to accept the potentially unwarranted assumption that a strictly linear model is adequate to describe the typical judge.

3. Significance tests are available to guide the nomothetic aggregation of ideographic descriptions of judges. Differences among the judges with respect to the weights they assign individual cues can be detected, and inferences can be made about cue utilization in the population represented by the judges.

4. Sources of unreliability that can be examined are extended beyond the standard G-theory approach and can inform the design of subsequent decision studies. For example, the impact of judges' differential cue utilization can be taken into account.

5. HLM analyses extend the lens model equation by permitting the decomposition of lack of agreement among judges into error variance and random effects of cues. This information subsequently can be used to guide training intended to improve interjudge agreement.

In summary, the HLM framework is a promising tool for combining SJT and G-theory into a single coherent framework that is simultaneously capable of assessing the dependability (reliability) of judgments and modeling diversity in judgment models across judges. In effect, G-theory is extended to take into account the information (cues) upon which judgments are based, and the lens model equation's pairwise consideration of interjudge agreement is supplemented by an analytic framework that considers all judges simultaneously, both individually and collectively.

Anyone desiring an electronic copy of the latest version of this manuscript may contact me at the e-mail address below.

Judgment Analysis, Think-Aloud Protocols, Cause Mapping, Image Theory and Neural Network Simulation

Ray Cooksey
Armidale, Australia

In the past year or so, I have been involved in four different areas of Brunswikian research with various colleagues and students at the University of New England, as well as in Europe.

The first research study involved a fairly standard judgment analysis approach to understanding job applicant shortlisting judgments made by a university job selection panel. Shortlisting here was defined as the decision to

allow an applicant to proceed to the interview stage of the selection process (otherwise the applicant is not considered further).

However, we coupled the standard judgment analysis methodology (using actual job applicants' curriculum vitae where the shortlisting outcome was known) with a think-aloud protocol methodology, applied on a judgment-by-judgment basis. This permitted us not only to model judgment policies using the cues established by university policy, but also to trace the potential dynamic intrusions of other information into the judgment process for each job applicant, both at the individual and group levels.

We are in the process of gathering data for a second judgment analysis investigation that also employs a think-aloud protocol method and qualitative cause mapping methodology. This time, we are examining university student judgments about engaging in risky sexual behavior (i.e., unprotected sex) using a series of representative scenarios (hypothetically generated using a broad-based survey of students to establish the 'population' parameters for sampling cue values and intercorrelations).

Part of this project is to get at students' subjective judgments of the risks associated with engaging in unprotected sex in a variety of different situations. Another goal is to tap into students' mental models regarding the factors they see as contributing to engaging in unprotected sex. Here is where we see the value of marrying judgment analysis, think-aloud protocols, and cause mapping methods in a coherent triangulated approach.

We are currently planning an investigation designed to provide insights into the dynamics of the commons dilemma, using a combination of judgment analysis methodology, image theory principles, and cause mapping methods. Here we plan to employ dynamic simulation software, specifically designed to run commons dilemma scenarios related to ocean fish harvesting, to look at fish harvesting judgments over time, made on the basis of exposure to key cue information about the resource pool and other external conditions and constraints. We also hope to map participants' mental models for the commons dilemma in terms of key factors that influence resource availability.

Finally, I am assisting a colleague, Hubert Bruins in the Department of Oral and Maxillofacial Surgery at the University Medical Center Utrecht, in his efforts to test a new methodology for judgment analysis. He has designed a neural network approach to judgment analysis that shows good potential for dealing with non-standard judgment tasks (such as those involving dichotomous judgments and categorical cues) where a multiple regression approach may be difficult to defend (i.e., in cases where assumptions are not met or where dynamic nonlinear relationships exist).

The specific judgment being modelled relates to the prophylactic extraction of teeth in patients with cancers of the neck and/or head prior to subjecting the patients to radiation therapy. The neural network approach models conditional probabilities as the basis for estimating cue-judgment and cue-cue relationships over a series of patient cases - the neural network is essentially grown and evolved as one progresses through judgments made on the patient cases. This study is currently being finalized in a paper to be submitted to Medical Decision Making.

Judgments about Union Representation, Crossing Picket Lines, Sexual Harassment Cases

Jim Holzworth
Storrs, CT

Steven Mellor, Jim Conway and I recently completed a manuscript concerning people's inclinations to be represented by labor unions. Employed persons (not currently members of labor unions) were asked to make judgments concerning how likely they would be to vote in favor of (or against) union representation.

Within-person analyses supported our hypotheses that policies would be influenced by perceived costs and benefits of representation. The direction of influence suggested that intent to vote for a union was lower when unions were perceived as being antagonistic, costly, exclusive, and corrupt, and higher when unions were perceived as providing a voice, a grievance process, a sense of security, and respect and dignity.

Between-person analyses supported our hypotheses that decision frame would be related to cost and benefit influences. Individuals who were asked to consider their vote as one in favor of a union indicated more intent to vote for representation. The corruption cue had a stronger negative influence on judgments when individuals were asked to consider their vote as one against a union.

Daniel O'Shea, Steven Mellor, David LaHuis, and I completed a study concerning contextual and individual influences on the individual's decision to become a replacement worker during a strike. Regression parameters from within-person analyses indicated that strike publicity, number of strikers, and threat of violence influenced individual judgment policies about willingness to cross a picket line to accept a position. Using these parameters as outcome variables, between-person analyses indicated stronger (negative) relationships between threat of violence and willingness to cross for those with low financial need than for those with high need.

Lisa Kath is analyzing her master's thesis data concerning judgments of sexual harassment federal court cases. We have judgment data from 53 jury eligible adults, 25 men and 28 women ranging in age from 18 to 78. Each person made a series of judgments concerning each of 50 cases described in narrative form, abstracted from court documents. We first compared judgments with court verdicts, using Swet's A statistic, and found an average A statistic value of 0.78. There was not a significant difference between men and women. We are now conducting judgment analyses to assess individual policies concerning different aspects of the sexual harassment judgments: severity of alleged actions, pervasiveness of alleged actions, sexual harassment verdict, and employer liability.

I am continuing my collection of biographical data (biodata) from university students in an attempt to relate it to styles of inductive reasoning. Cognitive Continuum Theory is guiding this research.

With encouragement from Ken Hammond, I am about to begin a classification of cue probability learning studies according to task characteristics, dependent variables, and anything else I can think of. I expect that this will take me a long time. An annotated bibliography of all published cue probability learning studies was prepared along with my contribution to The Essential Brunswik. The bibliography is available on the Brunswik Web Page:

Research in Patient Decision Support

Celia Wills
East Lansing, MI

I am currently engaged in two newly-funded projects conducting research in patient decision support. The first, "Information Interpretation in Patient Decision Support," funded by the Agency for Healthcare Research and Quality, is with PI Margaret Holmes-Rovner. The second, "Patient Decision Making About Antidepressant Medication," is a Clinical Scientist Career Development Award awarded by NIMH, for which I am the Principal Investigator.

Both projects focus on how patients evaluate information and make decisions relevant to their own health needs. The common goal of both projects is to develop knowledge for understanding how patient decision support interventions should be designed to foster more effective shared decision making between patients and health care providers.

The first project compares populations varied by ethnicity (African American/Caucasian) and education (with/without college education) to investigate patients' interpretation of information about risks and benefits of treatment decisions for benign prostatic hyperplasia (SDP/BPH) and the impact of such information on their decisions. We propose to examine information types commonly used to communicate risks and benefits of medical treatments to patients: statistical information, graphical lists, graphical drawings and diagrams, and patient interviews.

Research questions to be addressed are:

1) Do information types differentially affect participants' likelihood to choose a treatment?

2) Do information types differentially affect participants' interpretation of importance and salience of the information?

3) What new questions or concerns are raised for participants by full information?

The second project is an investigation of patients' decisions to decline or discontinue depression medication, despite its effectiveness and availability. Little is known about how people make depression treatment decisions, including key influences on decision making and appropriateness of decision making as related to health status and health system outcomes. Research on patient decision making can provide information that is needed to develop patient-focused interventions to improve depression treatment outcomes.

The goal of research in the first project is to describe relationships over time between patient decision making, medication use, health status, and cost and utilization of health services outcomes. Based on findings from the first project, a patient decision support intervention for primary care depression treatment will be pilot-tested for feasibility in the second project. The long-term goal is to improve the quality of primary care services for depression through implementation of decision support interventions for diverse populations of patients undergoing depression treatment.

Physician Decision Making in Testing for Prostate Cancer

Paul Sorum
Albany, NY

Why do so many primary care physicians routinely order prostate specific antigen (PSA) tests on their asymptomatic male patients over 50?

I have participated over the past year in a study of judgment and decision making about testing for prostate cancer in collaboration with Junseop Shim and Tom Stewart in Albany and with Gérard Chasseigne, his student (Sylvie Bonnin), and his physician (Joel Cogneau) in Tours, France. Junseop describes in his submission to the Brunswik newsletter the methodology of the study and his particular focus on physicians' perceptions of the guidelines.

Dr. Cogneau and I were primarily interested in trying to answer the question of why physicians routinely order PSA's in spite of the recommendations against this of their specialty groups. The study was designed to test seven hypothetical explanations: 1) that physicians who test routinely are not aware of the evidence-based specialty recommendations (but are likely to be aware of the position of the American Cancer Society); 2) that these physicians are deferring to patients' requests or demands (which are fueled by the media); 3) that they think that PSA tests provide very useful diagnostic information; 4) that they think that treatment of asymptomatic prostate cancer is more beneficial than the evidence suggests; 5) that they would greatly regret it if they did not order a PSA and the patient was subsequently found to have an advanced cancer; 6) that they fear that such a patient might sue them for malpractice; and 7) that they are uncomfortable with uncertainty (and hope to reduce uncertainty through testing).

We found evidence to support, to some degree, all of our hypotheses (with the exception of #6); the high rate of ordering PSAs is associated with a multiplicity of factors. Not yet tested is the hypothesis that physicians display consistency, in the sense that those who order more tests will tend to score more highly on all or most of the explanatory factors.

In comparing groups, French primary care physicians, on the whole, seemed to be less aggressive than American ones in trying to find prostate cancer in asymptomatic patients, and U.S. family practitioners less than general internists. These group differences were reflected in differences in the explanatory factors and are consistent with findings in other studies.

Scripts as an Alternative Framework for Describing Medical Decision Making

Robert Hamm
Oklahoma City, OK

This year we had a paper published in the volume on medical decision making in the Cambridge JDM series: Hamm, R.M., Scheid, D.C., Smith, W.R., and Tape, T.G. (2000). Opportunities for applying psychological theory to improve medical decision making: Two case histories. In G.B. Chapman and F. Sonnenberg (Eds.),

Decision Making in Health Care: Theory, Psychology, and Applications (pp. 386-421). New York: Cambridge University Press.

The first section reviews the various efforts that have been made to apply decision theoretic ideas or research products in order to improve medical practice. It describes projects that have applied decision analysis to individual patients, made evidence-based guidelines available to physicians, trained physicians to reason with analytical principles, provided decision-relevant information to physicians, or provided computerized decision aids.

The next section presents two interventions in detail, and analyzes the reasons they did not accomplish their expected goals. The first intervention was an attempt to make physicians reduce the proportion of sore throat patients to whom they prescribe antibiotics by training them to judge more accurately the probability each patient has a sore throat due to streptococcus. Their probability judgments become more accurate (and lower), but they still prescribed antibiotics to the same proportion of patients. Perhaps the assumption "Decisions are suboptimal because judgments are suboptimal" is not correct. Or maybe a judgment of diagnostic probability did not actually control the physicians' prescribing behavior.

The second intervention attempted to make physicians start talking about "end of life" issues with hospital patients and patients' families soon enough that the families could decide not to have "everything possible" done to revive the patient as he or she died. A central element of the study's intervention was to provide accurate (model based) estimates of the probability that the patient would die during this hospital admission. The intervention had essentially no effect on the use of cardiopulmonary resuscitation (CPR). The study assumed that if physicians were given information that decision theory says is relevant for decision making (survival probabilities with and without CPR, and patient preferences), they would use it. Perhaps the typical physician's end-of-life decision making strategy makes little use of such information.

The final section presents a descriptive metaphor: that physicians follow scripts when they make decisions about patients. We argue that interventions will be more successful if they are based on a clear understanding of the scripts physicians currently use, and if physicians are explicit about the alternative scripts that are proposed to improve their decision making.

For the Brunswik community, we would note that the systematic framework that we use for studying judgments or teaching judgment strategies (e.g., "What is the probability that a patient's sore throat is due to streptococcus?" as a function of N features) can as easily be applied for studying action strategies (e.g., "What would you do for this patient?" as a function of N features; or, "What is the probability that you would do X?" as a function of N features).

Physicians' Perceptions of Guidelines Regarding Prostate Cancer Testing

Junseop Shim
Albany, NY

My recent research, in collaboration with Tom Stewart and Paul Sorum, involved physicians' judgment and decision making about prostate cancer. We approached the issue from two different perspectives. One was the question, "Why do so many primary care physicians order prostatic specific antigen (PSA) tests for their

asymptomatic male patients in spite of the evidence-based recommendations against routine testing for prostate cancer?" which was presented at the 22nd Annual Meeting of the Society for Medical Decision Making. In this study, we compared U.S. physicians with French physicians. Paul Sorum describes this aspect of our study in his report.

The second perspective was the current debate over screening for the disease. We investigated the influence of the conflicting guidelines on physician judgment and decision making about screening and management of prostate cancer. More importantly, we explored the implications of these judgments and decisions for individual health care delivery and public health policy. The abstract of the paper is as follows:

The purpose of this study was to investigate primary care physicians' judgments and decision making about prostate cancer and to explore their implications for individual health care service and public health policy. Thirty-two primary care physicians from the Capital District of New York State participated in this study.

Judgment analysis was used to provide insight into the factors that physicians consider important in decision making about screening and management of prostate cancer. Two sets of 32 hypothetical patient scenarios were presented to the subjects. Cues in Set 1 were age, difficulty in urination, size of prostate, shape of prostate, and patient request for a prostate test. The physicians judged the probability of prostate cancer for each case and the likelihood that they would order a PSA test. In Set 2, PSA level replaced the patient request for a test as a cue. Physicians judged disease probability and likelihood that they would refer the patient to an urologist. A short questionnaire about knowledge of and attitude toward prostate cancer was also administered.

Physicians were classified as pro-screening, anti-screening, and conflicting according to their responses to the question, "What are the recommendations of official medical bodies about routine screening for cancer in asymptomatic males after age 50?" These recommendations have in fact been conflicting. Thirty-seven percent of the participants had inaccurate knowledge of guidelines, perceiving them as recommending either routine screening (28 percent) or no routine screening (9 percent).

While the pro-screening group was most likely to recommend a PSA test and to refer patients to an urologist, the anti-screening group was unlikely to recommend the test and referral.

The physicians in the conflicting group were highly polarized into two categories: They were either highly likely or highly unlikely to order a PSA test and to refer the patients to a urologist.

The results showed that physicians disagreed about the nature of current screening guidelines, and that their disagreement was reflected in the differences in PSA test ordering and patient referral decisions for a set of paper patients. In addition, physicians' disagreement was reflected in some patterns in their regular practice, and in their knowledge base regarding prostate cancer.

It was, of course, not possible to establish causality in this study, but there are important implications for individual health care delivery and public health policy that deserve further exploration.