THE BRUNSWIK SOCIETY
Newsletter

Click here for the Index

Table of Contents/Research Summaries

- [Lenard Dalgleish](), The University of Queensland
- [Patrik Juslin](), Uppsala University
- [Kathy Mosier](), NASA Ames Research Center
- [Reid Hastie](), University of Colorado
- [Werner W. Wittmann](), University of Mannheim

From the labs of John Gillis (GILLISJ@CLA.ORST.EDU) and Frank Bernieri (Toledo)

As some of you may recall, we have been busy applying the lens model to social judgments. Specifically, we are studying how observers estimate the level of social rapport that may exist between two people engaged either in a cooperative task or in an adversarial debate. We work with a video archive of 120 male and female college students who were videotaped in both cooperative interactions and debates with at least two different people (of the opposite sex) on different days. Subjects were unacquainted with each other. Interactions were self-terminated and lasted anywhere from 5 to 30 minutes. All interactants completed a 29-item rapport inventory that essentially measure how well they got along with each other (our rapport criterion). From this video archive we created sets of stimulus tapes that contained 37-50 brief video clips (30-50 seconds in length) that were extracted from each dyadic interaction. We then coded as many features as we could from these clips to generate a large array of potential cues to which judges might be responsive. We have been showing these stimulus tapes to judges, varying instructions and viewing conditions in an attempt to uncover the important aspects of their judgmental strategies.

After collecting more data, our first report has finally been published in the July issue of the *Journal of Personality and Social Psychology*. Results: (a) Ecology of self-reported rapport is encoded fairly well within a mere 50-sec slice of the behavioral stream. Cues (i.e., natural interaction behaviors) are nonorthogonal. (b) Ecology (behaviors that are diagnostic of rapport) differs slightly as social context changes from cooperative to adversarial. (c) Achievement (accuracy of rapport judgments or self-other correlations) is modest with $\underline{r}$'s ranging from .20 to .30. (d) Judgments of rapport rely on target expressivity (quantity, extremity, and variability) more than anything else. Judgment policies seem unresponsive to changing ecologies. Thus (e) social judgment was more accurate when observers were assessing rapport in the cooperative context where their judgment policies based on expressivity more closely matched the ecology (expressivity was more predictive of rapport, here).

This past year we have been examining the possibility that males and females differ in their responsivity to different kinds of training. We found, surprisingly, in previous studies that achievement for our task improved when outcome feedback was provided but did not improve when judges were told explicitly what cues they should use to maximize their accuracy. In this social judgment task, judges apparently were unable to intentionally employ a predefined policy. Accidentally, we uncovered a modest sex effect such that the above finding may be more true for females than for males. We repeated our feedback study this year and have focused our analyses on gender differences. Again, we have found that males appear to be more effective at implementing predefined rapport judgment policies than females, who still responded more favorably than males to outcome feedback.

Frank, particularly, has always been impressed by the fact that a 50-second slice of the behavioral stream taken out of a much larger interaction sequence contained so much valid information relating to a social outcome (i.e., rapport). With the help of Jon Grahe, a student of Frank's at Toledo, we have completed a series of conditions that pushed the limits of social judgment acuity by systematically reducing the stimulus length from the original 50-seconds to a mere 5-seconds. Different groups of judges viewed clips of 5, 10, 20, 30, and 50-seconds in length. Achievement varied across these conditions but not in a uniform (or even obvious) pattern.

We have also examined achievement as a function of information source (verbal versus nonverbal behavior). We isolated our information medium to different channels of communication including the pure Nonverbal (silent video screen display), pure Verbal (transcribed speech of the 50-second interaction segment), and Audio channels. Achievement in assessing rapport from our 50-second slice was highest in the pure Nonverbal behavior display condition and actually decreased to the extent that verbal information was made available to the judges. The two studies above suggest a lot about interpersonal processes and interpersonal perception.

We are continuing our cross-cultural replications of this work and now have data from Indonesia, Australia, Pakistan, Lebanon, Greece, and Mexico with additional samples being collected in Hong Kong, India, Russia, and Belgium. Of the 8 cultures analyzed so far, cross-cultural consistency in rapport judgment policies seem to be the rule. Both achievement and judgment policies from all countries are practically indistinguishable for our American samples. Although the Indonesian data was not different from the American data, the Indonesian Psychologists still maintain that the Indonesians are responsive to subtly different cue manifestations. Issue of cue definitions and cue coding were raised. All of this is intriguing given that the data collected showed no hint of what the Indonesians were suggesting.

We've also completed the first phase of our "Love Judgment" study. Couples living in Montreal were videotaped as part of a large couples study conducted by some clinical psychologists, David Zurroff and Richard Koestner. Among other things, couples completed several instruments that assess love for partner. We combined them all to create a composite measure of "love of partner." We assembled 50-sec video clips (extracted from 10 minute interactions) from the 48 couples on to a stimulus tape and showed them to judges in Toledo and Corvallis. We are analyzing that ecology now. Achievements (judge interactment and agreement) ranges from .10 to slightly over .2 depending on whether judges are measuring how much (a) the man is in love, (b) the woman is in love, and (c) the couple is in love. Achievement is highest with the female target, and lowest for the male target.

James Hogge (HOGGEJH@CTRVAX.VANDERBILT.EDU) and Stephen Schilling

In our 1995 JDM presentation, we demonstrated how the comprehensive analytic framework of generalizability theory provides the judgment analyst a convenient way to estimate the magnitude of multiple sources of error in order to design a study that can be expected to yield a required level of dependability (i.e., reliability). Such an analysis allows the investigator to examine the impact of averaging in various ways, e.g., across occasions, across cases, across judges (in apparent conflict with Brunswikian orthodoxy), or combinations of averaging. While such an analysis is very useful, it unfortunately ignores information about the cues available to the judges and thus tells us nothing about the extent to which variability among raters is due to policy differences.

We have also described the use of linear models that extend generalizability theory to include significance tests of possible nonlinearity of cue-judgment relationships and variability in judgment policies across judges. Now we are working on further development of this analytic framework in the hope of being able to use hierarchical linear models to provide explicit information about variation due to policy differences, thus paralleling the lens model equation's decomposition of observed correlations between pairs of raters.

Jeryl Mumpower (JERYL.MUMPOWER@CNSIBM.ALBANY.EDU)

I made the mistake of reviewing the paragraph that I wrote last year about "my recent Brunswikian-related research." I was not altogether pleased to find that I could submit more-or-less the same report this year. I guess I get good marks for perseverance and poor marks for creativity this term.

I have continued my work incorporating a Brunswikian-influenced view of negotiations. The paper by John Rohrbaugh and me entitled "Negotiation and Design: Supporting Resource Allocation Decisions through Analytical Mediation," which represents an effort to integrate and summarize our work during the last decade in this area, will appear soon in *Group Decision and Negotiation*.

Along with our colleagues Anna Vari and Tom Darling, John and I just completed and submitted a paper entitled, "Negotiation Support for Multi-Party Resource Allocation: Developing Recommendations for Decreasing Transportation-Related Air Pollution in Budapest." This paper describes our efforts to provide negotiation support for five task-force members who were trying to reach agreement about how to allocate a limited amount of money among programs intended to improve the air quality in Budapest.

Jim Sheffield, Tom Darling, Richard Milter and I completed and submitted a paper that reported on our multi-year, multi-study of interpersonal learning in negotiations. The gist of this work is that interpersonal learning in negotiations is not very good, but not because negotiators are afflicted by a "fixed-pie bias," as is widely claimed.

Along with Tom and Rick (co-authors on the previous paper), I have a paper coming out in *Acta Psychologica* that demonstrates that interpersonal learning is influenced by the interaction of substantive and formal task characteristics. (I think Brunswik would have approved.)

During the past year, I spent a lot of time thinking about trying to develop a computer-supported judgment aid for use in crisis decision making in psychiatric emergency rooms. As I reported last year, the decision is whether or not to admit people who present at psychiatric emergency rooms. The three most important factors are their mental status, the degree of danger they pose to themselves, and the degree of danger they pose to others. A number of imperfectly valid cues are associated with each of these factors. For the last few years, we have been in the process of trying to build judgment models that represent experts' judgments about the relations between cues and criteria. Presently, this work is stalled for lack of money.

Finally, after six years, Tom Stewart and I finally did succeed in completing our paper "Expert Judgment and Expert Disagreement" for the special issue of *Thinking and Reasoning* that Mike Doherty edited. As we worked in applied policy contexts, Tom and I realized that Social Judgment Theory is a powerful tool for understanding many expert disagreements, but that there are still a number of different sources of cognitive disagreement (leaving aside disagreement due to incompetence, venality, and ideology) for which we do not have good diagnostic or treatment tools. Tom and I discuss and categorize these sources of expert disagreement.

Ulrich Hoffrage (HOFFRAGE@MPIPF-MUENCHEN.MPG.DE)

1. Overconfidence and PMM Theory

In a series of experiments, I explored several aspects of PMM theory (Probabilistic Mental Models: A Brunswikian Theory of Confidence, Gigerenzer, Hoffrage, and Kleinboelting, 1991, Psychological Review). The main results were: Consistent with PMM theory, the hard-easy effect could be eliminated if both a hard and an easy item set were generated in a representative design. Results also showed that the PMM theory can make rather accurate predictions on the level of items. In addition, the size of reference class is an important variable,

both for over/underconfidence and the accuracy of those predictions. Finally, the results show little support for our use of PMM theory as an explanation for the confidence-frequency effect.

The manuscript about this research is on my things-to-do list. At the moment I can only send a copy of a poster that I presented at the last SPUDM (1995) in Jerusalem.

2. Bayesian Inference

Last year our article "How to Improve Bayesian Reasoning without Instruction: Frequency Formats" (Gigerenzer and Hoffrage, 1995, Psychological Review) was published. Although the title does not capture the relation to Brunswik, there is an important one: With "frequency formats" we refer to frequencies as obtained by natural sampling. As we stated on p. 686:

"Brunswik's 'representative sampling' is a special case of natural sampling." This term, natural sampling, which we borrowed from Gernot Kleiter means that there are no constraints on which observations will enter the sample and which will not. More recently, we were able to replicate the main result with physicians who worked on medical problems: Presenting the information in terms of frequencies as obtained by natural sampling considerably improves Bayesian reasoning.

3. Hindsight Bias

We (Ulrich Hoffrage, Ralph Hertwig, and Gerd Gigerenzer) developed a model that explains hindsight bias in terms of cognitive reconstruction: If people cannot retrieve their original judgment, they try to reconstruct the knowledge on which it was based, and from this reconstructed knowledge they infer what the judgment was. Because feedback is a cue that is strong enough to shift some knowledge states in the direction of feedback, this reconstruction of the original knowledge may not be veridical and, in turn, results in hindsight bias.

Translation for Brunswikians: the original judgment (before feedback) corresponds to "subject's response." After feedback, it becomes the distal variable in the memory task ("How did I answer this question last time?"). The feedback about the correct answer serves as a cue when an attempt is made to reconstruct the cue values of the original judgment. Thus, it causes systematic differences within the lens. We hope to submit the manuscript very soon.

4. Organization of Workshop: "BRUNSWIK TODAY: MODELS OF CUE-BASED INFERENCES"

Together with my colleague Ralph Hertwig, I organized a workshop at TeaP (one of the largest German psychology conferences). In the introductory talk, I gave an overview of Brunswik's life and work, and discussed the reception of his ideas in psychology. I suggested two topics in this workshop that would be of importance to Brunswikians: (a) the marriage between Brunswik and Darwin (captured by talks 1 and 2, see below), and (b) the exploration of fast and frugal decision algorithms as an alternative to standard statistical techniques such as multiple regression (captured by talks 3, 4, 5, 6, 9, and 10).

The following talks were given: 1. Cues for Mate Choice: Darwin, Brunswik, and Sexual Selection (Geoffrey F. Miller); 2. The Evolution of Cue Validities (Peter M. Todd); 3. Models of Bounded Rationality for Inference: Dealing with Constraints of Limited Time and Knowledge (Jean Czerlinski, Daniel Goldstein, and Gerd Gigerenzer); 4. Information-theoretical Aspects of "Take the Best" (Laura Martignon); 5. Hindsight Bias as a Result of Cue-based Reconstruction: Model and Simulations (Ralph Hertwig and Ulrich Hoffrage); 6. Hindsight Bias as a Result of Cue-based Reconstruction: Experimental Results (Ulrich Hoffrage and Ralph Hertwig); 7. Stereotypes as Distal Constructs in a Probabilistic Environment (Klaus Fiedler); 8. Effects of Perspective in Conditional Reasoning as Knowledge-based Inferences (Sighard Beller); 9. Cue-based Understanding of Conditional Statements (Alejandro López and Gerd Gigerenzer); 10. Decisions in Motion: An Approach to Dynamic Decision Making (Philip Blythe).

5. Information-theoretical Aspects of "Take the Best"

I recently started a cooperation with Laura Martignon (a mathematician) about information-theoretical aspects of "Take the Best" (this is the new name for the core algorithm of PMM theory). This cooperation is too young to provide results.

For further information (including my non-Brunswik-related research), please contact:

[Return to contents](#)

Gérard Chasseigne (chasseigne@univ-tours.fr)

Aging and Multiple Cue Probability Learning: The case of inverse relationships and irrelevant cues.

A large number of studies have shown that intellectual performance of the elderly is generally poorer than that of younger people. The performance of older people has been chiefly attributed to slower information processing. Reduced working memory capacity is one of the most notable consequences of this slowing. The effect of reduced working memory capacity, due to aging, on ability to learn direct and inverse probability relationships was evaluated in a Multiple Cue Probability Learning study using both outcome and cognitive (task information) feedback.

In a static forecasting task, 96 subjects equally divided among three age groups (20-30; 65-75; 76-90 year-olds), were asked to learn to predict the temperature of the water delivered by a cellar boiler. The elderly (65-90 years-olds) were not able to learn an inverse (negative linear) relationship with only outcome feedback.

In the same experiment, 48 subjects (16 in age each group) were given task information. The very elderly were not able to apply the knowledge of the inverse relationship provided by the task information. The less elderly did not find it as difficult to modify their cognitive functioning. Such modification nevertheless seemed possible only through an alteration of the task (counting rather than judging).

In a second study, 96 subjects participated in a static weather forecasting experiment. Elderly (65-90 year-olds) had difficulty learning probabilistic relationships, even direct relationships, when the task included irrelevant cues. Older people were not able to discount the irrelevant cues. Cluster analyses showed that the mean tendencies matched individual results.

84 subjects participated in a dynamic weather forecasting experiment. One cue lost its validity in the middle of the learning process. In the outcome feedback condition, elderly subjects were not able to learn to discount spontaneously the invalid cue. In the task information condition, when told that the cue would become invalid, older adults stopped utilizing the cue. Inhibition is a costly process, and this was evident in cognitive control. Although control was not affected by the utilization of inverse relationships by young people in our static experiments, it was clearly affected when discounting was required in a dynamic task.

A study is underway with the goal of evaluating the relative importance of cognitive aging factors in predicting knowledge in a linear and a non linear MCPL task.

Kim Vicente (benfica@mie.utoronto.ca)

Aside from my practical work in the area of cognitive engineering, I have been working on one particular project that might be of interest to Brunswikians. JoAnne Wang and I have written a paper (currently under review), entitled "An Ecological Theory of Expertise Effects in Memory Recall." The abstract is as follows:

Previous research has shown that there is a significant correlation between domain expertise and memory recall performance after a very brief exposure time. This finding has been replicated in a number of different domains, ranging from chess to figure skating, and is robust with respect to variations in methodology. Despite the large number of such studies (over 40 in total), there are a number of findings in the literature for which there is no satisfactory theoretical explanation. As a result, the boundary conditions under which expertise effects are to be expected have yet to be clearly ascertained. In this paper, a novel product theory based on ecological theories of skill acquisition is proposed to explain existing findings on expertise effects in memory recall. The theory, referred to as the constraint attunement hypothesis, provides a framework for identifying and representing the various levels of goal-relevant constraint in a domain. Given this description of the environment, the theory predicts that there will be a memory expertise advantage in cases where experts are attuned to the goal-relevant constraints in the material to be recalled, and that the more constraint available, the greater the expertise advantage can be. The proposed theory explains a number of diverse empirical findings in the literature in a coherent, unique, and parsimonious fashion, and suggests a number of promising issues for future research.

Robert Gifford (snap@uvvm.uvic.ca)

Here in Victoria, my research group continues to use standard and adapted lens models to explore and compare perceptions and evaluations of and by various groups:

1. Perceived and measured intelligence, with nonverbal and verbal behavior cues measured during interviews. The judges are unacquainted peers.

2. Attractiveness of voice-ad placers, with vocal cues measured from their voice ads. The judges were unacquainted listeners.

3. Competence of applicants for teaching jobs, with verbal and nonverbal cues measured during actual job interviews. The judges are school principals.

4. Vulnerability of single-family dwellings to burglary, with physical features of the houses as cues and burglars, police, and residents as judges.

The write-ups of these studies have been delayed by the completion of the second edition of my textbook on Environmental Psychology, which is being printed right now and will be available about September 1, from Allyn and Bacon. This shameless plug may be justified by saying that Brunswik and the lens model receive star treatment in the introductory and perception chapters of the book.

Ray Cooksey (rcooksey@metz.une.edu)

Beyond Judgment Analysis: Brunswik and Systems Theory

I completed my book on "Judgment Analysis" for Academic Press in July 1995, and it appeared in print in January 1996. I am now in "arm chair" mode and have moved on to do some critical and integrative thinking about what Brunswik's ideas have to say about broader perspectives regarding human behavior. In particular, I am exploring the links between Brunswikian theory and method and systems-based approaches to understanding human decision behavior in complex naturalistic contexts. Some of my thinking is building upon an early unpublished CRJP paper written by Mumpower, Adelman, and Rohrbaugh in 1975, where they attempted to integrate Brunswik's ideas with those of the systems theorist, C. West Churchman. I want to take these linkages further, if possible, to create a unified systems-based perspective which describes human decision behavior. In line with Ken Hammond's recent discussion of the differences between coherence-oriented theories and correspondence-oriented theories, I am looking to construct a correspondence-based perspective which can integrate theory, method, and context. Initially, the rich context I want to explore is managerial decision making in organizations. Recent systems views which have emerged in organizational research have focused on the learning organization, the utility of metaphors and other idiographic phenomenological methods for gaining insights into personal perspectives of organizations, and on the applicability of chaos/complexity theory as a way of capturing the nonlinear dynamics of organizational systems. These perspectives have some clear implications for how managers make strategic, as well as routine, decisions and how they learn from their mistakes. Brunswik, of course, also had important things to say about how people will operate and learn within an ecological framework. It is time these varied perspectives were brought together. Two years ago, I described an extension to the Judgment Analysis method which incorporated concepts from fuzzy set theory and chaos theory (this extension was subsequently incorporated into the last chapter of my book). However, I feel there is much more to be gained by taking Brunswik's ideas beyond the traditional boundaries of Judgment Analysis. Once I have derived the integrated perspective, I will then move to empirically examine its implications.

Return to contents

Claudia Gonzalez-Vallejo (GONZALEZ@OAK.CATS.OHIOU.EDU)

Since last November, we have made great progress in our medical decision making study on the management of acute otitis media (AOM) in children. This is a collaborative project between Albany Medical Center and the Center for Policy Research. Using Judgment Analysis as a framework, we looked at the factors that influence physicians' diagnosis of AOM and their treatment choices. Thirty-two physicians in the Albany area judged the presence of AOM and made treatment decisions for 32 hypothetical patient cases. Each patient case contained information on historical and examination variables that are typical of patients seen by physicians. Variables included the age of the child, the frequency of past AOM, temperature, and the degree of redness, bulging, and mobility of the tympanic membrane. For each case, physicians were asked to provide a probability judgment of the presence of AOM and to choose whether to treat with antibiotics or not. If antibiotics were prescribed, they were further asked to select a specific antibiotic and to select reasons for their choices. Two sets of results emerged from our study: a high degree of uniformity across physicians' judgments of the presence of AOM and variable management strategies. First, physicians' probability estimates were based primarily on examination variables and, in particular, on the variables describing the ear drum status. Multiple regression analyses predicting each doctor's probability judgments from the patient information showed that physicians' judgments were well described by a linear combination of the case variables; R^2s ranged from 0.76 to 0.98 with a median equal to 0.93. We also observed a high level of agreement among physicians' judgments as depicted by high judgment intercorrelations. On the other hand, the minimum probability level resulting in the prescription of

antibiotics (i.e., a threshold) varied from physician to physician, as well as their antibiotic prescription patterns. Furthermore, different patient cases yielded more or less judgment agreement. Cases producing higher variability of judgments across physicians may be considered ambiguous and they resulted in more treatment disagreement than non-ambiguous cases. More physicians prescribed antibiotics for ambiguous cases when the mean judged probability of AOM was low rather than high. Overall, the level of judged probability of AOM, the ambiguity status of the case, and the degree of bulging and lack of mobility of the ear drum of the case were significant predictors of antibiotic usage.

In the weather forecasting arena, Tom Stewart and I have continued thinking about a naturalistic weather forecasting study where we can bring together calibration, lens model, and signal detection theory notions. So far, we have looked at new analyses of forecasting data previously collected in Colorado and Albany; the data contain rain forecasts for each of four judges and hail forecasts for each of seven judges. In these two contexts, we know, approximately, what information forecasters used to form their forecasts and so we know the relationship of available cues and the actual outcomes. From Brier type analyses, we saw that both calibration and resolution were better in predicting rain than hail, resulting in better skill for rain forecasters. Conditionalizing the data on the event and looking at response distributions, we see that discriminating rain from no-rain is much better than discriminating hail and no-hail (d's were much greater in the rain than the hail situation). Furthermore, in both situations we predicted judgments very well from the information used by the forecasters. In the case of rain, however, available cues predict rain quite well; in the case of hail, they do not. These two results, in part, explain why skill is poorer in predicting rain than predicting hail; hail forecasters, nevertheless, could improve skill if they reduced both the mean and the variability of their responses.

Return to contents

Ken Hammond (KHAMMOND@CLIPR.COLORADO.EDU)

With the exception of a regrettable illness that prevented me from attending the 1995 Brunswik Society meetings, the past 12 months were good ones for me. I finished my book manuscript titled "Human Judgment and Social Policy: Irreducible uncertainty, inevitable error, unavoidable injustice" and managed (with the invaluable assistance of Mary Luhring) to get it into print by the time I am writing this. It took me about four or five years to get all this material in intelligible form (here I must give credit to my editor at Oxford University Press, Joan Bossert, who kept demanding intelligible form). The final product was somewhat surprising to me. I will be grateful for your comments.

During all this time I had the strange feeling that I had been working on this manuscript for all of my professional life, yet was uncovering new ideas as I worked out old ones. I didn't do justice to many of these, but I was greatly comforted by the knowledge that Ray Cooksey, that solid citizen of the Brunswik Society, was bringing out his book, "Judgment Analysis: Theory, Method and Applications" (Academic Press), and, therefore, the reader of my book who wanted to know how all these ideas got worked out ("Is there any beer in all this foam?") could discover that in Ray's highly readable book. It was a pity that we couldn't have published our two books at the same time with the same publisher (and matching covers?), but Ray was about nine months ahead of me. Be that as it may, anyone interested in Brunswikian ideas can now find their basic mathematical substance in Ray's book, and the manner in which these ideas can be used to explore the formation of social policy in mine. In all immodesty, I must say I am very happy indeed that these two books are now available. At last we have a direct and up-to-date answer to those who want to know where they can find (and taste both beer and foam in) neo-Brunswikian psychology.

One new idea that is developed in my book is that of the role of evolution in Brunswikian psychology. Although Brunswik makes clear his Darwinian foundations, and many of us have mentioned them over the years, this aspect of Brunswikian psychology has never been explicated. I tried to remedy this omission as I incorporated

and made use of the two major metatheories of coherence and correspondence to unify the field of judgment and decision making. It will be interesting to see how these ideas are received. Perhaps I can present their essence at the '96 meeting and thus obtain some criticism. (Mike Doherty: Will you be at the meeting?)

I am now about half way through a book ms on: "Judgment Under Stress." I will try to develop further the theory presented in "Human Judgment and Social Policy" and apply it to this important topic. I have now done enough work to see that this will be a fruitful enterprise, and that the Brunswikian approach will again turn out to be a very useful guide to a new area of application. I will appreciate receiving from members of the Society any information that they think might be useful to me in this enterprise. Rob Mahan has already conducted some very interesting and important empirical research based on the application of cognitive continuum theory to stressful work. A request for reprints sent to RISK@UGA.CC.UGA.EDU will, I am confident, be productive.

[Return to contents](#)

Rob Hamm (rob-hamm@uokhsc.edu)

Robert M. Hamm, Christie Gilbert, and Mary Lawler.
Investigation of Teens' Strategies for Avoiding Violent Relationships.

Work in progress:

Judgment analysis techniques were used to investigate male and female teenagers' perception of the possibility of violence in their relationships. Each subject (27 females, 23 males) performed three different judgment tasks: judging date attractiveness, probability of violence at a party, and whether someone should terminate a relationship.

For example, Task 2 described a party at a boy's/girl's house, where the host puts a sexual demand on a girlfriend/boyfriend. Three cues were varied in the stories -- whether the host's same-sex parent had ever hit the spouse, whether there is alcohol at the party, and whether the host has been violent with the guest in the past. There were three levels of the third cue (no violence, threw keys, hit him/her). The task was to judge the chances that the host will hit the guest the night of the party or in the near future. The response scale required the subject to circle one of the numbers 0%, 2%, 5%, 10%, 20%, 35%, 50%, 65%, 80%, 90%, 95%, 98%, 100%.

Each task was done twice, once from the perspective of one's own sex (for Task 2, for example, the guest at the party was the same sex as the subject), and once from the perspective of the opposite sex.

Analysis will include (1) comparison of lens model (subject side) features such as control (R-squared) and relative weight on each dimension; (2) comparison of males with females; (3) assessment of subjects' conception of the opposite sex -- do they think the opposite sex is just like them?; (4) assessment of accuracy of subjects' conception of opposite sex -- does the average male judge the average female accurately? (5) Comparison of weights upon situational aspects (e.g., effect of date's history of violence upon date's attractiveness) with demographic and experience reported on a separate instrument (e.g., one's own experience with violence in relationships).
[Return to contents](#)

Mike Doherty (mdoher2@bgnet.bgsu.edu)

There are a few things to talk about from northwest corner of Ohio.

1. I was asked to edit a special issue of the journal Thinking and Reasoning devoted to Social Judgement Theory. There was marvelous cooperation from all contributors, and it is due out as a double issue in August. Here is the Table of Contents:

Editorial

by Michael E. Doherty and Jonathan St. B. T. Evans

Social Judgement Theory

by Michael E. Doherty and Elke Kurz

The Methodology of Social Judgement Theory

by Ray W. Cooksey

Social Judgement Theory and Medical Judgement

by Robert S. Wigton

Expert Judgement and Expert Disagreement

by Jeryl L. Mumpower and Thomas R. Stewart

Self-insight, Other-insight and Their Relation to Interpersonal Conflict

by Barbara A. Reilly

Man as a Stabiliser of Systems: From Static Snapshots of Judgment Processes to Dynamic Decision Making

by Berndt Brehmer

Upon Reflection

by Kenneth Hammond

The contributors are widely representative of ages, continents, and areas in the field. The issue has history and theory, methodology, reviews of applications, a report of an experiment, and a marvelous retrospective by Ken. Note that I have not defected from American spellings; *Thinking and Reasoning* is published by Erlbaum in England. I am personally delighted with the product, and I see it as a vehicle for bringing two research traditions together.

2. Greg Brake, Gernot Kleiter and I have completed two calibration studies, both with Gigerenzer's, Bjorkman's, and Juslin's arguments about representativeness of sampling and with Erev, Wallsten and Budescu's paper on reliability and regression to the mean in mind. One was with the typical almanac questions. The other used baseball games with expert (a la Arkes' test) subjects, and with the locus of the task uncertainty also in mind. That is, we had the data for both alternatives squarely in front of the subject. The almanac questions yielded horrendously low reliabilities and the appearance of overconfidence. The baseball data showed a high degree of

UNDERconfidence for 17 of 20 subjects, and overconfidence for only one subject, with two being well calibrated. This was submitted to JDM, and we hope to submit it to either OBHDP or JBDM soon.

3. I will be giving a plenary address at the Third Annual Conference on Thinking in London in August. In that address, I will use Jungermann's Two Camps on Rationality as a jumping off point to discuss Cognitive Continuum Theory and the Six Modes of Inquiry, but I will be making an argument that might not sit well with some, and that is that much irrationality comes from a mismatch of the task and the strategy that may have been induced by the framing of the task.

4. I fear that the other lines of research I am pursuing would not qualify as Brunswikian, except in a very deep sense that we are looking for the same phenomenon across a variety of tasks, but that phenomenon is one that leads to poor performance rather than high achievement. A paper that should come out in the next issue of Memory and Cognition will explore the features of two apparently similar probabilistic tasks that lead in one version to excellent performance, but in the other to terrible performance, in both instances using a coherence criterion.

5. Another project just completed involved sending essays written by advanced graduate students about their lives to practicing clinicians. The study can be considered one on framing, or confirmation bias, or hypothesis testing.

The between Ss manipulation was asking them to say whether the writer was

1. psychologically healthy (Yes or No)

2. showing psychopathology (Yes or No)

3. psychologically healthy or unhealthy

We expected the subjects to be biased by the questions so that the same essays would be seen as

1. healthy

2. pathological

3. mixed responses because the essay data would be related to both hypotheses

Of course we got the opposite results

1. mixed

2. healthy

3. mixed

We speculated that the reason was that clinicians had a higher threshold for inferring "psychopathology," so we ran another condition:

4. psychologically unhealthy (Yes or No)

Conditions 1,3 and 4 did not differ at all.

How is this at all in the tradition of Brunswik? We sent out 9 separate essays, only one to each clinician in the assigned group. (We sent out over 400, got about a third back, all Ss were members of 2 APA divisions). If we had sent out just one essay, we might have concluded just about anything, depending on the one sent out; the variation in response to individual essays for two of them ranged from 100% healthy to over 90% unhealthy for one of them. Three of the 9 had a 50/50 split. This paper is written and if someone would like to see it, let me know.

[Return to contents](#)

Tom Tape (tgtape@mail.unmc.edu)

Bob Wigton and I have continued our work testing the effect of various types of feedback on students' ability to distinguish bacterial from viral meningitis. Since this is a complex, multivariable task with both negative and non-linear cues, we would have expected task information and cognitive feedback to be most effective. We reported last year at the J/DM meeting that in fact probabilistic feedback was most effective in simulated cases. We have now analyzed students' performance on paper cases abstracted from actual patient cases.

The results, which will be presented at the October Medical Decision Making Meeting, are similar to the simulations in that only probabilistic feedback significantly improved performance. Task information and cognitive feedback appeared to affect diagnostic decisions and antibiotic prescribing decisions by changing the decision criterion on the ROC curve without improving overall diagnostic accuracy. These results continue to support the hypothesis that feedback of calculated probabilities from a decision rule is a powerful learning tool.

[Return to contents](#)

Lubomir Kostron (KOST@PHIL.MUNI.CZ)

This is a report on activities at the Masaryk University, Brno, Czech Republic during the last year. As you may recall, we have done some translations of papers written within the Brunswikian framework by some of you. These translations have been used in my teaching of a one-semester course concerning Brunswik and later developments. Bright students made translations available with e-mail addresses of authors to students over INTERNET, so students from our other universities may download them. A set of these translations have also been mailed to the Psychology departments at the Charles University in Prague and Palacky University in Olomouc. The Psychological Department of the Czech Academy of Science also obtained a full set.

There was a M.A. thesis completed and defended, which used POLICY and CHAOS programs to study how judgments over fictitious admission examinations data will look. Ray Cooksey helped us a lot, providing data for us to use while learning the software. Really, it was learning and methodological study, showing us possibilities of both approaches, without interesting conclusive results. The author, Martin Vaculik (who is now in the PhD program) will give the summary of his thesis below. He turned out to be very bright, independent and clever. He learned much by himself. We need to get him on a study stay abroad, to learn with people who directly work with POLICY and are good statisticians. The summary of Martin's thesis follows now:

"Linear and Non-Linear Decision-Making: Experimental Comparison"

by Martin Vaculik

The subject of my paper is decision-making processes and the use of computer analysis in the investigation (mapping) of these processes. The classical approach of judgment analysis, represented by the computer program POLICY PC 3.0, and the approach based upon the chaos theory, represented by the CHAOS SYSTEMS SOFTWARE 6.1 program, were applied. The paper consists of a theoretical part and an experimental part.

The relationship between the decision-making process and other mental processes is described in the first chapter of the theoretical part. In this chapter, there is also a general description of the decision-making process, the development of decision-making methods, and the types of decision-making processes and strategies.

The probabilistic functionalism and its place in the field of psychology is discussed in the next chapter. The principles and the methodology of probabilistic functionalism and its use in the decision-making process are described, and possibilities and limits of Brunswik's approach are considered in the end of the theoretical part. The aim of my research was to provide a description of the decision-making process by means of the above-mentioned methods, and a presentation of advantages and limits of these approaches. The results were presented in both numeric and graphic form.

There were two types of data samples (artificial data and real data) processed by various methods. The first method was the standard graphic presentation; the second one was the computer program POLICY PC 3.0, made exclusively for the judgment analysis; and the third way of data processing was the CHAOS SYSTEMS SOFTWARE 6.0 program, which had been originally created for the analysis of time series of economical data, but it could also be applied in the decision-making processes.

The different types of data samples were processed. One data sample was gained in a planned situation, the other data were simulated. The artificial data were used to demonstrate results obtained within exactly defined conditions. To simplify the presentation of results and to make it more clear, the typical cases of judgment values (low or high correlation between judgment and cues, non-linear relationship between judgment and cues) were used.

The results of the research demonstrate possibilities of both programs in the analysis of decision-making processes. The CHAOS SYSTEMS SOFTWARE 6.0 program can be used to complete results gained by the POLICY PC 3.0 program. No data have been processed in different time points (just one occasion, when judges met). The analysis of this kind of data could be a subject of further research.

Martin also demonstrated "group conferencing" with POLICY to some staff and faculty members - discussing the issue, how a new library of the school should look. We had some interesting views expressed, and the method was introduced to the school administration.

My own work, stimulated by a recent visit with Prof. Hammond in Boulder and the SUNY at Albany group, culminated in my current writing of a volume, trying to present the Czech audience with a readable:

- summary of Brunswik's probabilistic functionalism and representative design,

- an overview of later developments (I have traces of), including SJT and the theory of cognitive continuum,

- an analysis of what remained avoided or unsolved (which, of course, is the most difficult part, considering the base I can work with and my own limited knowledge)

- an outlook for the future and connections I see to some other approaches outside of Brunswikian tradition proper (modeling of dynamic systems, information theory, some play with extensions of the lens model into various formally different situations).

What I missed, while meeting some of you, was more examples of strategic thinking and perhaps accommodations of new developments outside of the scope of your usual everyday thinking - I refer to your regular meetings at the Brunswik Society. Could you devote the next one to this? I hope to bring some input of this kind in writing.

The major change, however, is that I am not alone any more, and a second generation of Brunswikians is appearing even here. I wish to express my gratitude to Prof. Hammond, John Rohrbaugh, Tom Stewart, David Andersen and to Ray Cooksey for continuous support. Since I am useless in empirical research and very limited in knowledge of statistics, would you kindly try to see if there are chances to make possible proper experiences directly to our students?

This brings me to concluding thought - plans for the future. My plan is to continue with translations (Brunswik is so rich, as I read him again now), and some papers by Hammond and Brehmer are still to be translated to complete the basic set. Also, there will be a seminar devoted to the use of the POLICY program so more people will learn how to handle it and see its applicability without being scared off by Brunswik's difficult stuff. Then there will be more of us.

Rob Mahan (RMAHAN@EGON.PSY.UGA.EDU)

After what seems to be an eternity of grant writing endeavors, we have now configured a modest laboratory here at UGA dedicated to the study of judgment and decision making activities within the Brunswikian framework. Of particular interest to us have been the effects of work-stress on the judgment process. We have focused on time-on-task and uncertainty manipulations and have found some interesting time-on-task decrements associated primarily with consistency in the execution of judgment protocols. In addition, we have recorded time-on-task by uncertainty interactions that appear to suggest a sort of threshold that demarcates a shift in processing strategies. For example, we have found that time-on-task seems to mediate a shift in strategy from analysis to intuitive processing (i.e., over time people begin to manifest intuitive processing during the performance on an analytical task). Task uncertainty appears to increase the speed at which this change in processing strategy occurs during a continuous performance judgment scenario.

At the moment, we are extending the above ideas using a distributed hierarchical approach where a team leader must organize information from subordinates who are tasked with sharing information about a criterion state in forming their expert judgments. These judgments, which represent different aspects of expertise concerning the criterion are passed to the leader for assimilation into a global judgment of the criterion. Preliminary data suggests that, within this continuous performance paradigm, consistency decrements are occurring at the subordinate (expert) level. Suprisingly, we are seeing matching decrements at the team leader level in the absence of consistency problems.

Finally, we are also examining some information display issues within the above context (i.e., information representation and work-stress). Continuous and sustained decision making operations are becoming more important in many business, government and military domains. It is often difficult, we believe, to make wholesale generalizations using past judgment research to settings that are highlighted by people working continuously for extended periods of time. One of our long range interests is in developing technological countermeasures (employing display design techniques) against decrements in judgment and decision making

that occur as a result of work-stress factors. These countermeasures would enhance, we believe, the behavioral and psychopharamcological efforts being developed primarily by the military.

We hope to share some notes with interested people at the coming meeting. Believe us, we need all the help we can get!

Jim Holzworth (HOLZ@UCONNVM.UCONN.EDU)

Janet Barnes-Farrell and I continue our judgment research concerning performance appraisal. We're still analyzing and writing up some results of earlier research. We have also begun an investigation of judgments concerning accommodation of employees with disabilities. The Americans with Disabilities Act (1990) will undoubtedly have an impact on the workplace. Along with a graduate student, Chris Daniels, we are designing a study of co-workers' judgments of fairness (social and organizational justice) of various accommodations for disabled workers. Some cues that will be manipulated are: nature of the disability, origin of the disability, safety issues, various mitigating factors, personal characteristics, performance level, organizational policies and practices.

At the 1995 Brunswik meeting, I had the pleasure of meeting and visiting with Celia Wills, a professor of nursing at Michigan State University. She has revived my interest in nursing judgment and decision making. We have recently submitted a manuscript to *Research in Nursing and Health* concerning judgments to seclude and/or restrain psychiatric patients. Nurses made judgments of the likelihood that they would recommend seclusion, restraint, or other nursing interventions for patient profiles containing 17 cues relevant to clinical judgments about use of seclusion and restraint. Cues about current behavior (clinical status) of profiled patients had the most impact on nurses' recommendations for use of seclusion and restraint, and nurses had good insight into the form of their own judgment policies. Nurses generated similar strategies for alternative nursing interventions to seclusion and restraint for the profiled patients. However, no one patient received identical recommendations for interventions from all nurses, and, overall, nurses agreed with each other on specific recommendations only about a third of the time. The results of this study show that nurses were in general agreement about types of strategies which would be used for management of psychiatric care situations. However, results raise the possibility that experienced nurses may differ substantially from each other in their clinical judgments about the specific nursing actions which should be taken for a specific patient care situation. The lack of agreement among nurses has implications for development of staff training programs which emphasize use of critical clinical and contextual cues, use of decision thresholds, and creative alternatives to use of seclusion and restraint to manage problematic behaviors in psychiatric care settings.

I will be happy to discuss these studies at our annual meeting, November, in Chicago.

Len Adelman (LADELMAN@GMUVAX.GMU.EDU)

Much of my effort this past year has gone into preparing a text entitled, "Handbook for Evaluating Knowledge-Based Systems: Conceptual Framework and Compendium of Methods," with Sharon Riedel of the Army Research Institute. It has nine chapters: (1) Introduction, (2) Literature Review, (3) Conceptual Framework, (4) Requirements Validation, (5) Knowledge Base (KB) Validation, (6) KB Verification, (7) Usability Evaluation, (8) Performance Evaluation, and (9) Planning and Managing the Evaluation Process. As you might guess, we

use the lens model as our conceptual representation for how evaluators should use test cases to evaluate the quality of a knowledge base. If all goes well, the handbook will be published by Kluwer Academic Publishers by early next year.

Terry Connolly (connolly@ccit.arizona.edu) and Adrianne Casebeer

Health care evaluators make extensive use of an instrument known as the SF-36, which is thought to measure various aspects of one's overall health. The questionnaire asks 36 simple health-related questions ("Does your health limit you in lifting or carrying groceries?"; "How much bodily pain have you had in the past 4 weeks?"); a number of simple index scores are computed from the responses.

We have been concerned that these health assessments are a function not only of the respondent's symptoms but also of their medical knowledge (accurate or not). That is, in an extension of our cues and components work, the overall assessment is determined not just by the cue values, but by the meaning the respondent attributes to these values. For example, knowledgeable patients may feel worse, or better, than less knowledgeable patients with exactly the same set of symptoms--diagnosis, or admission to medical school, can make you feel worse! In an experiment, we asked subjects to assess the overall health of 30 paper patients described by SF-36-type symptoms, after being instructed in the symptoms, treatment and prognosis for a fictitious disease. As predicted, the instructions led to systematic inflation or deflation of overall health ratings for appropriate paper patients. Policy capturing allowed us to explore the underlying cognitive processes of the subjects. The findings seem to have important policy implications, given the widespread use of SF-36 in health policy research. They also carry some theoretical implications in the cues/components ideas I have presented at earlier Brunswikian meetings.

Mark Chaput de Saintonge (D.M.ChaputdeSaintonge@mds.gmw.ac.uk) and Roy Poses with Wally Smith, Tom Tape, Bob Wigton, Constantine Gatsonis, Joachim Krueger, and Steve Sloman.

"Why do doctors differ when the evidence is clear? Studies of medical practice variation in the use of ACE inhibitors in heart failure."

Over the past few years, we have looked at the causes of practice variation between physicians using essentially non-Brunswikian methods. Our interest has centered on the investigation and management of cardiovascular disease in general and ischaemic heart disease and heart failure in particular. These diseases, their investigation and management are all associated with high morbidity and cost, and are the subject of high levels of public and health service concern.

We have recently started a series of studies to probe the causes of medical practice variation in this area and examine some reasons why physicians make treatment choices which do not appear to be based on the best available evidence.

Our study has three major objectives:

First, to establish how well calibrated are physicians' judgments about treatment effects compared to published evidence.

Second, to explain why deviations may occur even in well-calibrated physicians by testing a series of structured hypotheses addressing specific cognitive biases. For example, belief that new and "high-tech" treatments produce better outcomes may lead to their over-utilization and relative underutilization of older and "low-tech" alternatives.

Third, to explain why unfavorable practice patterns seem so resistant to most interventions. Any intervention that does not speak to the processes underlying physicians' decisions is unlikely to work. If a physician avoids a drug because he or she judges that the drug frequently produces specific adverse events, interventions that simply remind the physician about the drug or provide physicians with data about therapeutic outcome probabilities are likely to fail.

The study will be largely vignette-based. More general clinical scenarios will be used to elicit judgments on outcomes of treatments in populations of patients as distinct from individuals. These will be compared with those available in published criterial outcome studies. The design we propose will allow us to distinguish the operation of cognitive biases from a simple failure to recall the published data. The results should suggest how to design interventions which target the specific problems physicians have judging particular outcome probabilities that lead to unfavorable practice patterns. These interventions could be developed and evaluated in future studies.

The first study will look at the use of ACE inhibitors in heart failure and will examine the judgments of UK family practitioners, internists and cardiologists.

This will specifically examine the use of these drugs in heart failure of various causes. Patient vignettes with a very general frame will probe respondents anticipations of benefit and adverse drug effects in populations satisfying the entry criteria to major trials in mild-moderate heart failure. We anticipate that our three groups will show systematic differences here. Specific and systematically constructed vignettes will examine whether prescribing decisions are inappropriately influenced by cues which resemble predictors of adverse and favorable outcomes but are not themselves predictors. We will also be able to see whether newer (more expensive) ACE inhibitors are preferred for the more severe case and older drugs for the less severe. Other drug treatments will feature among the cues, and this should give us some idea of respondents' views of the place of ACE in the prescribing hierarchy.

Return to contents

Peter Juslin (peter.juslin@psyk.uu.se)

In recent research, we have been concerned with a distinction between two different origins of uncertainty in judgment, which we refer to as Brunswikian and Thurstonian. Brunswikian uncertainty reflects states of limited knowledge due to merely correlative properties between known and unknown attributes of the world. Thurstonian uncertainty refers to stochastic variation and imperfections of the information processing system itself, and is thus a property of the organism. We believe that there are some interesting differences between the psychological processes that are relevant to these two origins of uncertainty, and these differences surface most clearly in connection with perceptual judgments. We have recently developed a computational model relevant to one class of perceptual tasks dominated by Thurstonian uncertainty, sensory discrimination with the method of pair-comparisons. The sensory sampling model is a sequential sampling model that predicts decisions, calibration of subjective probabilities, and the complex pattern of response times in simple psychophysical discrimination. The distinction between the two origins of uncertainty and the sensory sampling model are

presented in a forthcoming *Psych. Rev.* paper (Juslin & H. Olsson, in press). We are currently exploring and testing the model, with a particular eye to the relationships between the model and the concepts of detection time (mental speed) and the possibility of extending the model to more complex perceptual judgments.

A recent trend in research on overconfidence initiated by Erev, Wallsten and Budescu, and others is to stress the importance of random error in the probability judgment process. In a recent paper co-authored by one of the "early" Brunswikians, Mats Björkman, (Juslin, H. Olsson & Björkman, submitted) we explore the consequences of adding different assumptions about the random error component (a Thurstonian component) to the Brunswikian framework provided by Gigerenzer's PMM-theory. One interpretation of the error component is as sampling error in experience and/or memory retrieval; another interpretation is in terms of response error in the process of overt probability assessment. These two interpretations both predict conditions of bias with hard-easy effects, but nevertheless lead to different patterns of predictions. For example, the assumption of response error leads to prediction format-dependence, where we can get different and even contradictory results when we use different probability assessment scales (i.e., underconfidence in half range assessment with simultaneous overconfidence with full range assessment). A (still distant) goal of our future research on probability assessment in cognitive tasks is to integrate theory on probability judgment with cognitive theory on memory and categorization, with a particular interest in the exemplar-based algorithms that respond to both frequency and similarity (representativeness) in determining classification decisions.

Other current interests in our lab include the development of a theory on hindsight phenomena that can predict powerful (confirmed) reverse hindsight phenomena (Winman, Björkman & Juslin, in preparation), the application of calibration techniques for investigating the confidence-accuracy relation in eyewitness identification (e.g., Juslin, N. Olsson, & Winman, JEP: L, M, & C, forthcoming September issue), and base rate effects in categorization.

Return to contents

Tom Malloy (malloy@uconnvm.uconn.edu)

Linda Albright and Thomas Malloy just completed a paper (now under review) titled "Integration of the Approaches of Brunswik, Campbell, and Cronbach in Research on Interpersonal Perception." This paper shows that a common conceptual and analytic thread can be found in Brunswik's theoretical and analytic work, Campbell's mutitrait-multimethod matrix, and Cronbach's generalizability theory that is integrated and extended by the Social Relations Model developed by David Kenny. Albright and Malloy have collaborated with Kenny in a number of studies of interpersonal perception, dealing with both substantive phenomena (consensus, accuracy, self-other agreement, cross-cultural perception, consensus across social contexts) and analytic issues focusing primarily on componential analysis. Work in progress focuses on specification and estimation of structural equation models using maximum likelihood estimation procedures to estimate the relationship of multiple stimulus cues to perception, judgment, and behavior. Both mediated and unmediated models are being considered to broaden the application of such analyses to different theoretical approaches (e.g., Gibsonian and Brunswikian).

Return to contents

Marilyn Rothert (ROTHERT@msu.edu)

The MSU Decision Making group continues to work with judgment/decision making regarding menopause. After completing our intervention study, we replicated the judgment study with low income African American

Women. This study asked the question "How do women use information to make decisions regarding hormone replacement therapy at menopause?" Date were collected from 200 women using the cases from our previous study updated and modified for the population. Analyses indicated few policies that were significant. Secondary analyses of the judgments revealed lack of variability in judgments, a high likelihood that these women would use HRT, and use of the factor "hot flashes" almost exclusively. We are trying to further analyze these findings to assess the potential for future application of judgment studies to lower SES populations.

We submitted a competing continuation proposal based on our previous work extended to low income African American women. The proposal was submitted July 1 to NIH, National Institute of Nursing Research. The team continues to publish from these studies. Members of the team include Marilyn Rothert, Margaret Holmes-Rovner, David Rovner, Celia Wills, Georgia Padonu, Geraldine Talarczyk, Jill Kroll, and Neal Schmitt.

[Return to contents](#)

Frank Laufer (fnl01@health.state.ny.us)

Can decision conferencing ensure fairness in group decision making? Previous research has demonstrated the effectiveness of decision conferencing using social judgment analysis software such as POLICY-PC in the context of a group decision making process (McCartt and Rohrbaugh, 1989). However, also of interest is whether individuals participating in or observing the process perceive the conduct of the conference and the judgment policy it produces to be fair or just. Six criteria have been proposed by Leventhal (1980) as procedural rules by which people can judge the procedural justice of an allocation process - consistency across people and time, bias suppression, accuracy, representativeness, correctability, and ethicality. The applicability of these procedural criteria to individuals' perceptions of the fairness in hiring decisions, performance evaluations, and other interactions with legal and managerial authorities has been demonstrated. John Rohrbaugh and I are using a survey which incorporates items intended to measure each of these criteria and the overall fairness of the decision conference process and the judgment policy it produces. Surveys have been sent to and returned by individuals who have either participated in or observed three conferences conducted during the past two years. We are constructing scales for each of Leventhal's procedural rules from the survey responses and will examine their correlations with peoples' overall judgments of procedural fairness of the decision conference process.

[Return to contents](#)

David Funder (FUNDER@UCRAC1.UCR.EDU)

The "Riverside Accuracy Project" continues its examination of the moderators and processes of accurate judgment of personality attributes. I published my theoretical formulation concerning the accurate judgment of personality, called the "Realistic Accuracy Model" (or, regrettably, RAM), last September in Psychological Review. Readers of this list will readily recognize it as a close relative of Brunswick's' lens model. We are mapping empirically demonstrated moderators of accuracy onto this model, and continuing to analyze a very large data set concerning the personalities and judgments of personality of a sample of about 180 UC Riverside undergraduates.

[Return to contents](#)

Alex Wearing (ajw@ariel.ucs.unimelb.EDU.AU)

Our (Mary Omodei and I) work this year has focused on two matters. The first is adapting our Firechief simulation to study distributed dynamic decision making situations. In essence, that means linking several PCs together, with each PC having access to partial information about the terrain. Information and resources can be transferred between "commanders", depending on the structures and rules imposed by the experimenter. This simulation is a kissing cousin of Berndt Brehmer's work, and perhaps others of whom we are unaware. Josh Klayman is also significantly involved in this project.

The second matter has been the development (particularly by Mary Omodei and Jim McLennan) of a light head mounted video camera to facilitate gathering information about real time decision making by e.g. firefighters, sports referees etc.

Apart from that, two further activities. Josh Klayman and I have been making slow progress on a study, reported in part by Josh last year) of the ability of subjects to detect the structure of simple dynamic systems. Second, Oswald Huber of Fribourg in Switzerland spent a semester in Melbourne in the first half of the year, and we have almost completed an empirical paper on the simulation of subjects' strategies in multistage decision tasks, focusing on Oswald's Breeding Lizards task.

I plan at this stage to be in Chicago in November, by which time we should have some preliminary results from the distributed ddm and the head mounted video tasks, and there may be even a beta version of the software.

[Return to contents](#)

Alex Kirlik (kirlik@isye.gatech.edu)

We've written two things that may be of interest to some of the Brunswikians. The June issue of *Human Factors* saw publication of our paper "Supporting perception in the service of dynamic decision making" (with Neff Walker, Dan Fisk, and Karin Nagel). This paper describes two studies we did to measure whether we could support performance in dynamic tasks by training and design interventions targeted toward fostering perceptual and pattern recognitional processes. The second of these studies evaluated the "ecological task analysis" methodology I discussed at a past Brunswik meeting. Empirical results are encouraging.

The second thing is a chapter I recently wrote for a forthcoming book called "A Companion to Cognitive Science" edited by Bill Bechtel and George Graham. From the authors notes this is targeted to provide an encompassing overview to cognitive science, and is organized in a number of general issue areas including cognitive activities, methodologies, theoretical stances, controversies, and "cognitive science in the real world." For this latter area I was asked to write a chapter on "The design of everyday life environments." The main point of the paper is to chronicle how design and studies of humans interacting with artifacts in everyday situations has forced a rethinking of some central cognitive science issues. The reason I mention this work here is that the essence of the paper provides evidence and argument for the need for theory of the functional system comprised of the human and environment considered collectively. The paper mentions that the empirical data in the field is prompting an increasing number of researchers to reconsider earlier functional or ecological theorists. Brunswik is mentioned along with Vygotsky and Gibson in this regard.

If you would like a copy of either of the above please email me. I should mention that the second paper is still with the editors.

[Return to contents](#)

Tom Stewart (T.STEWART@ALBANY.EDU)

While working with Paul Roebber and Lance Bosart on a set of data from students and faculty forecasting precipitation and temperature in a natural setting, I learned, again, how important it is to pay attention to the task when analyzing expert judgment. Accuracy was very high for all forecasters, as was agreement (Jim Shanteau tells me that their agreement may be the highest ever seen in an expert judgment study). Differences between forecasters were small relative to differences between tasks, and task differences could be predicted from an understanding of critical task properties (such as number of cues, cue intercorrleations and, most importantly, task predictability). The abstract for the paper, which has been in review for an exceptionally long time, appears below.

Stewart, T. R., Roebber, P. J. and Bosart, L. F. (1996). The importance of the task in analyzing expert judgment

The accuracy of judgmental forecasts of temperature and precipitation was analyzed. In contrast to the findings of many studies of expert judgment and forecasting, forecasts were highly accurate and forecaster agreement was high. Human forecasters performed better than an operational forecasting model and about the same as a linear regression model. Differences between temperature and precipitation forecasts could be predicted from differences between the tasks. In general, differences between tasks were much greater than differences between forecasters. Task predictability was an excellent indicator of forecast accuracy. The characteristics of the environment for forecasting temperature and precipitation that contribute to accuracy include high-quality information and the availability of "guidance" based on a computer model. It is concluded that an understanding of the properties of the task is essential for understanding the accuracy of expert judgment.

Return to contents

Joshua Klayman (fac_jlk@gsbvax.uchicago.edu)

I'm not up to anything new this year, but I am working on two ongoing projects. One is a set of studies of overconfidence done with Jack Soll and Claudia Gonzalez-Vallejo. We thought we were through with this months ago, but the more we look at our data, the curiouser they get, and new ideas about how to analyze confidence-judgment data are published almost weekly, so this has become a surprisingly long-running project. (Thanks to Jon Leland at NSF for being understanding about extending the time frame on our grant.) In particular, we are wrestling with the question of what is a "real" hard/easy effect, and how would you know it if you had it.

The second ongoing project is in collaboration with Alex Wearing at the University of Melbourne, and was started while I was visiting there last year. We are interested in how people learn aspects of dynamic systems, such as chains and cycles of causation. We have data now on how people learn some of the elementary particles of such systems, i.e., various causal links among three variables with random error. One tentative finding is that people find it difficult to learn when the same variable is both a cause and an effect.

Return to contents

Len Dalgleish (len@psy.uq.oz.au)

Over the last decade or so I have been developing a general risk assessment and decision making (RADM) model. The model has been applied in a number of diverse situations, e.g., Child protection risk assessments

and decisions; the decision to use protective equipment in risky workplaces; ethical decisions by students; and in auditing, the decision to qualify a set of accounts (more details later).

Very briefly, the RADM model is based on the idea that in decision making in risky situations the assessment of the amount of risk in the situation is separated from the decision about the acceptability of the risk. That is, for the decision maker there is a threshold of acceptable risk such that if the assessment of the amount of risk in a situation exceeds it, action is taken because the risk is unacceptable. There are two parts to the model. One part models the risk assessment. An overall risk assessment is thought of as comprised of at least three components: magnitude of future harm, chances of future harm, and strengths of situation. Assessments of these risk components is based on the situation characteristics. For each decision maker, a hierarchical lens model (judgment analysis) is used to estimate the paths from situation characteristics to risk components and from the risk components to overall risk. The other part models the placement of the threshold for acceptable risk. The factors that influence the threshold come from the decision maker's history and experience, and not from the characteristics of the current situation. This part of the model is based on a signal detection theory analysis of the task. The task for the decision maker is to detect the need to take action to avoid harm, that is to discriminate those situations in which future harm will occur if no action taken from those in which future harm will not happen.

The RADM model has some testable predictions (thankfully) and in my current research it is being tested in the laboratory using a version of Berndt Brehmer's and Leif Loveborg's simulated firefighting task, NEWFIRE, as the task. A lab task is needed because to test the model, a situation is needed where the researcher knows the true outcome of each situation being judged. In all the other situations where the model has been applied, the correctness of the decision cannot be assessed because there is no gold standard, e.g, in child protection decision making.

Over the last few years, my students in their fourth year Honours thesis project have used a variety of tasks to demonstrate that the model can describe the assessment of the amount of risk in a situation and the placement of the threshold for acceptable risk. A partial list is:

1. Young people's decision to overtake a car on a bend.

2. Young people's decision to accept a ride with an intoxicated driver. Do high risk takers perceive the situations as having a lower amount of risk or do they have different thresholds for acceptable risk? (The answer is a bit of both.)

3. Auditors', partners in one of the big accounting firms, decisions to qualify a set of accounts. They differed more in their thresholds for acceptable risk than in the assessment of the amount of risk.

4. Students' decisions to cheat in an assignment (give their assignment to another student). Students with different types of ethical orientations differed in their thresholds for acceptable risk, but not in the assessment of the amount of risk in a situation.

5. Small business owners' decisions to expand their business.

6. Industrial worker's decisions to use personal protection, e.g, ear muffs, in hazardous situations. (Project is currently underway by my PhD student, Heidi Bushell.)

The model and my own research on child protection decision making, e.g., the decision to make a case a child protection matter, the decision to separate the child from the family, and the decision to reunify child and family, is being incorporated into a book. The book *Risk and Decisions in Child Protection* is under contract with Wiley(UK) and is due for publication in early 1998. The manuscript is due in August 1997.

Congratulations to Ken Hammond on his book finally being available. I got a flyer in the mail about the book yesterday. I am looking forward to reading it since the applications of my RADM model are in situations of great uncertainty with great costs associated with the inevitable errors.

Return to contents

Patrik Juslin (Patrik.Juslin@psyk.uu.se)

Brunswikian Music Psychology

I became familiar with Brunswik's ideas through the works of Mats Bjorkman, Berndt Brehmer, and Peter Juslin. Surprisingly, I realized that Brunswik's metatheoretical framework could be adapted to the study of music performance.

I have recently conducted some research on emotional communication in music performance. More specifically, I have studied how performers communicate representations of basic emotions to listeners by means of a number of probabilistic cues in the performance, e.g., tempo, dynamics, timbre, articulation, and timing. For theoretical guidance, I have turned to a functionalist perspective which involves the integration of ideas from psychological research on emotion and nonverbal communication with Brunswik's probabilistic functionalism.

I have used a modified lens model to conceptualize the communicative chain which consists of (a) the performer's expressive intention, (b) the expressive cues in the performance, and (c) the listener's judgment of the expression. Multiple regression is employed to capture various aspects of the communicative process. The tools of judgment analysis described by Cooksey (1996) are really extremely useful. Reports on this work will be published soon.

Return to contents

Kathy Mosier (mosier@eos.arc.nasa.gov or kmosier@mail.arc.nasa.gov)

We (Linda Skitka at UIC and I) have been continuing our research on the use of automated cues in decision making situations. Last year at the meeting, I gave a preliminary report on a part-task aviation simulation study that I conducted at NASA Ames, demonstrating that pilots tended to rely on heavily automated cues and that the presence of these cues short-circuited the information gathering process. We are getting ready to run a follow-up on that study, using a two-person (crew) configuration, and focusing on interventions that will foster the cross-checking of automated information with other available cues. Linda has been doing parallel studies in her university lab. We have found similar patterns of automation use and errors across populations, and will be investigating the effectiveness of similar interventions. At JDM this year, I will be presenting the final reports on the first studies at UIC and NASA (we won't be finished with the two-person studies yet).

The research emphasis in aviation is shifting toward Free Flight issues - i.e., issues in the plan to introduce more shared control of the airspace (e.g., giving air crews more responsibility for choosing their own flight paths and maintaining their own separation). One of the issues we will be investigating has to do with the relationship between the possession of information (i.e., cues) and responsibility for decision making. What are the implications, for example, of air crews possessing traffic information that previously was only available to controllers? If anyone is aware of literature that ties responsibility/blame for decisions with knowledge of information or cues, I would be glad to hear about it.

Reid Hastie (reid.hastie@colorado.edu)

Dan Gigone (now at the Fuqua School, Duke University) and I have continued our research on the "common knowledge effect" in group decision making. Essentially, the more members of a group who know a decision-relevant cue, the greater the impact of that cue on the group judgment or choice. A new paper on a choice task (before we've always studied estimation tasks) is in press in the Journal of Personality and Social Psychology and a paper on analyzing judgment accuracy (which includes a Brunswik-inspired decomposition of sources of accuracy and error) is in press in *Psychological Bulletin*.

David Rettinger (Psychology Department, University of Colorado) and I are finishing a paper on an application of lens model policy capturing methods to infer which types of cues (e.g., necessity, abnormality-in-a-background-situation, etc.) people rely on when inferring the strength of causal relations among concrete singular causal events presented in typical newspaper story reports. Subjects make judgments of the causal strength between candidate causes and an effect that is the subject of each of several news articles. Regression models are used to index the weight of the abstract relationships (e.g., necessity) that might be the basis for the assessment of subjective causal strength. There was considerable between subject variation in the nature of the causal strength assessment policies; some of the differences appear to derive from individual differences in emphasis on statistical regularity cues versus explanatory mechanism cues.

Alan Sanfey (Psychology Department, University of Colorado) and I have just completed a project that replicates and extends a study by MacGregor and Slovic of reading graphic displays that communicate the data (runners' training, motivation, age, and past performance) for judgments of marathon finishing times. We found, for this task, that text displays, compared to bar graphs and tables, increased the weight placed on cues expressing the runners' motivation and produced the highest levels of predictive accuracy (ra).

Werner W. Wittmann (www@tnt.psychologie.uni-mannheim.de)

The Mannheim research program on Brunswik-symmetry

Brunswik-Symmetry and the lens model equation

Principles of symmetry are the key concepts of all successful true sciences. Detecting or conceptualizing something important for winning a Nobel prize in science is correlated with applications of principles of symmetry. The laureates Richard P. Feynman and Murray Gell-Mann are prototypical examples in physics. Brunswik's representative design and the related lens model incorporate all of the virtues and promises of symmetry. Tucker (1964) provided us with a mathematical equation (the lens model equation) which can be used to test hypotheses under what conditions predictions and explanations succeed or fail. Tuckers equation explains a given observed correlation as a function of a linear and a nonlinear term. Each term is multiplied and thus attenuated by the linear and nonlinear models of predictor (PR) and criterion (CR) respectively. We have modified and enriched that equation with a random error component, the psychometric reliabilities of PR and CR, and a selection effect parameter mapping restriction or enhancement of range.

We have applied principles of Brunswik-symmetry to many different areas of scientific psychology with great success as regards prediction and explanation, e.g. meta-analysis of psychotherapy outcome research,

personality research, attitude research, intelligence and working memory capacity research, intelligence and complex problem solving and program evaluation research. Some of these papers are in english some in german and not easily accessible to an international audience but I will bring some of them with me to Chicago. The majority of these conceptualizations, ideas and relationships to reliability, validity and generalizability can be found in Wittmann (1988).

Wittmann, W.W. (1988). Multivariate reliability theory: Principles of symmetry and successful validation strategies. In Nesselroade, J.R. & Cattell, R.B. (Eds.). Handbook of multivariate experimental psychology. 2nd ed. (pp. 505-560), New York: Plenum Press

For more detailed information, including the modified formulation of the lens model equation, contact:

Prof.Dr.Werner W. Wittmann
Lehrstuhl Psychologie II
Universitaet Mannheim,Schloss ,EO
D-68131 Mannheim,Germany

Voice: 49-621-292-5639
Secretary:49-621-292-5640
Fax:49-621-292-2528
Email :www@tnt.psychologie.uni-mannheim.de

Return to contents

Index