The Brunswik Society

Newsletter

Vol. 27, 2012

Edited by Lars Sjödahl, lars.e.sjodahl@gmail.com

Many thanks to all authors for their contributions

To get information about future Newsletters,
you are welcome to join (free of charge) the email list of the Brunswik Society at:

http://www.brunswik.org/

Table of Contents

EDITORIAL

Sjödahl, L…….. 4

OBITUARY of BERNHARD J. WOLF

Sjödahl, L. ……..…….…….... 5

CONTRIBUTIONS

Afonso, J., Garganta, J., McRobert, A., Williams, A. M., & Mesquita, I.
The Perceptual Cognitive Processes Underpinning Skilled Performance in Volleyball: Evidence from Eye-Movements and Verbal Reports of Thinking Involving an In Situ Representative Task……..…… 8

Athanasou, J. A.
Judgment and the Probability of Responding…….……... 11

Cavagnaro, D. R., Gonzalez, R., Myung, J. I., & Pitt, M. A.
Adaptive Design Optimization for Risky Choice Experiments…… 13

Dhami, M. K.
Heuristics and Visual Aids…… 15

Fiorentini, C., Mortillaro, M., Bänziger, T., & Scherer, K. R.
Studying Emotional Communication from a Brunswikian Perspective……..… 18

Hamm, R. M.
Two Methodological Variants on the Lens Model…… 20

Hammond, K. R.
Different Meanings of the Concept Representativeness…….…..… 22

Holzworth, J.
News from the Past Year…..... 24

Hong, S.-K.
Judgment Gap Analysis between Service Provider and Consumer for Service Design……. 25

Jiang, Z. & Nagasawa, S.
A Model Study on Emotional Communication in a Mono-Brand Fashion Store – Application of the Lens Model in the Fashion Industry…… 26

Kaufmann, E.
Meta-Analysis Research under the Critical Lens: A Missing Brunswik Link?……... 27

Mitchell, G.
The External Validity of Research in the Psychological Laboratory……... 29

Mumpower, J. L.
News from the Past Year……..... 31

Nestler, S. & Back, M. D.
Applications and Extensions of Lens Model Analyses in Personality Psychology……... 32

Pachur, T., Hertwig, R., & Steinmann, F.
How Do People Judge Risks: Availability Heuristic, Affect Heuristic, or Both?..... 34

Persson, J. & Wallin, A.
Why Internal Validity is not Prior to External Validity: Paper presented at PSA 2012…… 37

Rolison, J. J.
Dual-Processes in Learning and Judgment:
Evidence from the Multiple Cue Probability Learning Paradigm……. 39

Sjödahl, L.
Lose your Forest and You Lose Everything……... 40

Stewart, T. R.
Judgment Research at Albany…….. 45

Wecker, C.
Evidence-Based Practice in Education Requires (Quasi-)Representative Designs……. 46

Weidman, A. C. & Holtzman, N. S.
A Brunswikian Lens Model of Personality Impressions from Facebook Profiles…….. 49

Wittmann, W. W.
Brunswik-Symmetry……... 52

NOTICES

Marewski, J. N. & Hoffrage, U.
Call for Papers……. 53

NEW BOOK

Sjödahl, L.
Psychiatric Diagnoses – Are they to be Trusted? First Listen to the Patient!……..... 54

Editorial

Lars Sjödahl
University of Lund,
Sweden

Contact: lars.e.sjodahl@gmail.com

Today's social mobility and information exchange between different social and cultural contexts have increased demands on individuals' adaption to trans-cultural values such as openness, mutual understanding, ability to cooperate, respect and empathy in interpersonal relations. These demands require compromises, adjustment to norms, to ambiguity and uncertainty as well as awareness of consequences, corresponding to our decisions and choices. Brunswik's conceptual world, theory and lens model paradigm are well-suited for empirical studies of these adaptive processes.

The contributions to this year's Brunswik Society Newsletter cover a wide range of topics: judgements when no beforehand, agreed upon criterion exists, emotional communication, sport, crime, medical risk communication, just to mention some examples. On the theoretical side, the validity concept is discussed and Brunswik's representativeness-concept illustrated against the background of the subjective probability concept.

During the last 6 years I have been the editor of the annual Brunswik Society Newsletter. I have learnt a lot, inspired by the correspondence and always enjoyed the contacts with colleagues. I also wish to extend my very grateful thanks to Esther Kaufmann who with her considerable knowledge of Brunswikian research has directed my attention to interesting work and presumptive contributors. She has, moreover, taken on the responsibility for the layout of the Newsletter. I am also indebted to my wife, Gillian, for IT and language support. Due to advanced age, family commitments and a multitude of other rather time-consuming interests I have decided to make this issue my swansong. The Brunswik Society has kept up a long, unbroken flow of information about Brunswikian research by means of our annual Newsletter. I sincerely hope that one or some of you will take up the torch and keep our Newsletter going.

Obituary

Bernhard J. Wolf

Lars Sjödahl
University of Lund,
Sweden

Contact: lars.e.sjodahl@gmail.com

Bernhard J. Wolf, Brunswikian psychologist, professor and researcher, died on March 3, 2012 in Landau. He is survived by his wife, Gunhild, and son, Jonas.

Bernhard was born on April 26, 1948 in Bonn, where he studied psychology at the university, 1966-1971. From 1971 he worked as a psychologist and educational scientist at the University of Koblenz-Landau. His fields of research and teaching were wide, encompassing: socio-scientific research and statistics, pedagogic psychology, development psychology, pedagogics of infancy, Egon Brunswik's work. Bernhard was well attuned to the scientific language of Brunswik who aimed at precision when explaining complex ideas.

In Bernhard's Habilitation work, "Brunswik und Ökologische Perspectiven in der Psychologie", (1993, published 1995) Brunswik's key-concepts and theoretical framework are presented against a background of manifold ways of describing and modeling human beings' environments. Referring to psychologists of that time like Kaminsky (1978) and Kruse (1978), traditional psychology is characterized by a general neglect of the environment. This disregard still abides for a number of reasons, one of which may be scarcity of interdisciplinary projects. Besides a variety of ecological model-illustrations Bernhard presents a comprehensive survey of psychologists, who in different ways are related to Brunswik's conceptual world.

Early in Bernhard's writings we find an emphasis on structural variations of research perspectives. Brunswik, familiar with visual perception studies of objects, strongly recommended variations of research perspectives in psychology, for example, combining quantitative with qualitative data. This complementing is inbuilt into the circularity of Brunswik's lens model, i.e. feed-back arches from achievement to the subject. Such task feedback has often to be given in qualitative terms in order to function as reinforcements in learning situations. This perspective variation, exemplified in Bernhard's longitudinal studies of schoolchildren's motivational development (1997), is still an urgent issue in psychological research.

In a comprehensive survey over ecological perspectives in psychology during the past century, Bernhard illustrates Bronfenbrenner's four-level, hierarchical, social-cultural model with an interaction flow (Wolf, 1995, p. 205) where the macro-level is represented by a state (a nation) and micro-level by single families or individuals. The macro-system is not seen only as an organization but mainly as a trustee of laws, recommendations and rules based on single members' ideologies and values as expressed in daily-life and cultural activities. To fulfill these obligations the macro-system is greatly dependent on feed-back information, i.e. first-person accounts from single individuals representing the "lower-level" functioning, i.e., meso- and micro-

systems. Bernhard illustrates this feed-back dependency with a model, where the micro level is represented by families, single family-members and neighbours. This information flow (problem) has grown in actuality and complexity with time, due to increased demands on the individual's adaption to transcultural values.

In two longitudinal studies (1997, 6 years, and 2001, 10 years), Bernhard illustrates how Brunswik's request for representative design is applied when studying school children's "persistence", a motivational state defined as: "Persistent behaviour means to be able to remember an unfinished goal and to find solutions to complete it. Especially in children persistent behavior influences their learning with respect to performance, because persistence fosters reaching one's goals." (Wolf, 2008, p. 52). The actual situation sample is based on the following main information sources: study supervisors, kindergarten teachers, parents and elementary school teachers. Ratings were collected, containing 26 concrete situation descriptions. It was thus possible to follow the single child's development longitudinally. Comparisons were made between children with opposing persistence-profiles. Likewise, also high-persistence children were compared with regard to their situation ratings. These comparisons formed a basis for Bernhard's discussion about alternative generalizations over time.

Bernhard's initiative to arrange the "International Meeting on The Original Brunswik", in Landau, 2008 was greatly appreciated. To this meeting he called leading researchers from the US as well as from places around Europe and Australia. The program started with two lectures introducing the participants to the intellectual, trans-disciplinary environment, (the Vienna Circle), in which Egon Brunswik wrote his dissertation thesis and post-doctoral work. The meeting was held in a very cooperative and stimulating atmosphere, much due to Bernhard's planning and the comprehensive agenda. Among the many interesting themes only a few can be mentioned on these pages, for example, single case research, intuition and complementary ratiocination, measurement and uncertainty, environmental complexity.

My wife and I first met Bernhard at the Brunswik Society meeting in Minneapolis in November 2004, but it was not until 2008 that we really learnt to know him. He was then planning the "Original Brunswik" meeting and decided to come to Lund for a few days to discuss with me various subjects around Brunswik and his philosophy. It turned out that Bernhard had been to Lund before and knew his way around. He spent many hours at our house and we took him out to see local places of interest to him. He particularly enjoyed our visit to the ancient church in Dalby outside Lund. We learnt to know Bernhard during these days in Lund as a friend and colleague with good will, humanity and an exceptional sense for essentials in Brunswik's writings.

Despite many years of poor health Bernhard kept up his research and his contacts with colleagues. He was a loyal contributor to this Newsletter.

Bernhard was a generous and inspiring friend and our thoughts go to his wife Gunhild and his son Jonas.

Some important references of Bernhard´s work:

Wolf, B. (1995). *Brunswik und ökologische Perspektiven in der Psychologie*.[Brunswik and the ecological perspective in psychology]. Weinheim: Deutscher Studien Verlag.

Wolf, B. (2001). *Origins of the basic behavior-principles Univocality and Equivocality in Brunswik's system: Fritz Heider, Edward C. Tolman, Egon Brunswik*. Paper presented at the 17th Meeting of the Brunswik Society, Orlando, Florida, USA.

Wolf, B. (2004). *Representative longitudinal design and persistence. Empirical results in Educational Psychology.* Paper presented at the 20th Meeting of the Brunswik Society, Minneapolis, Minnesota, USA.

Wolf, B. (2005). A new type of person profiles by the Representative Design. In C. Harries (Ed.), *The Brunswik Society Newsletter, 20,* 20-21. Retrieved from http://www.brunswik.org

Wolf, B. (2008). *Quantitative single case research based on the concept of Representative Design introduced by Brunswik (1956).* Paper presented at the Swiss Federal Institute of Technology Zurich, Zurich, Switzerland.

Wolf, B. (2008). *Principles of Brunswik's probabilistic functionalism*. Landau, Germany: Verlag Empirische Pädagogik.

Contributions

The Perceptual Cognitive Processes Underpinning Skilled Performance in Volleyball: Evidence from Eye-Movements and Verbal Reports of Thinking Involving an In Situ Representative Task

José Afonso & Julio Garganta
University of Porto, Faculty of Sport, Portugal

Allistair McRobert & Andrew M. Williams
Liverpool John Moores University, School of Sport and Exercise Sciences, UK

Isabel Mesquita
University of Porto, Faculty of Sport, Portugal

Contact: jneves@fade.up.pt, jafonsovolei@hotmail.com

An extensive body of work has focused on the processes underpinning perceptual-cognitive expertise (Dicks, Button, & Davids, 2010; Ericsson & Williams, 2007; McPherson & Kernodle, 2007). In sports, particular attention should be allocated to the processes that underlie expert performance (Ericsson, 2008; McRobert et al., 2009). Due to the highly specific nature of expert performance (Williams et al., 2008), this aim should be accomplished using tasks demanding that athletes produce perceptual responses and motor responses, therefore respecting the perception-action couplings (Gibson, 1979). This will make research findings more generalizable for practice, as they will more closely resemble live-action scenarios (Dicks et al., 2010). In light of these considerations, we consider it highly relevant to design representative tasks, approaching the ecology of real-life performance (Brunswik, 1955) and the true nature of the processes underpinning expert performance (Mann, Abernethy, & Farrow, 2010). In order to achieve this goal, thoughtful consideration should be given to task nature and complexity (McRobert et al., 2011) and experimental design (Button et al., 2011; Dicks et al., 2010), respecting task specificity (Ericsson & Ward, 2007).

However, the majority of research published in this field has made use of film-based simulations to capture expert performance. In such cases, due to different perception-action couplings, the generalizability of results may be compromised (Araújo, Davids, & Passos, 2007). In our study, we combined eye movement recording and verbal reports of thinking to explore the processes underpinning skilled performance in a complex, representative volleyball defensive task involving in situ data collection. Altogether, 27 female volleyball players (15 skilled and 12 less-skilled) performed as center backcourt defenders in training sessions (simulated 6 vs. 6 playing) while wearing an eye-tracking device (Applied Science Laboratories 3000 MobileEyeTM registration system; Bedford, MA, US), and had to attempt ball interception when defending. After each sequence, athletes were questioned concerning their perception of the situation. The protocol of McPherson (2000),

adapted to volleyball by Moreno et al. (2008) and Araújo, Afonso and Mesquita (2011), was used. Reliability of the observation was tested for 22.2% of the data (randomly selected), using an intra-observer coding and an inter-observer coding. Cohen's Kappa ranging from 0.807 to 1.000 for categorical variables, while Cronbach's Alpha ranged from 0.900 to 0.977 for continuous variables. The level of significance was 0.05.

Our results showed that the visual search strategies employed by the skilled players were more exploratory than those used by the less-skilled players, involving more fixations ($F_1 = 4.792$, $p = 0.030$, $\eta^2_p = 0.029$) to a greater number of locations ($F_1 = 4.238$, $p = 0.041$, $\eta^2_p = 0.026$), in line with the results of North et al. (2009). The skilled participants had superior mean fixation duration, but without statistical significance. With regard to fixation location, the concept of functional visual space was defined: this concerns visual locations that are intermediate to two or more relevant visual cues. For example, that player may gaze upon a location equally distant to the setter, the ball, and the middle-attacker, using peripheral vision to capture information from several locations simultaneously (Behrmann & Ewell, 2003; Laurent & Ripoll, 2009; Williams et al., 2011). Results revealed that skilled participants spent more time fixating on functional spaces between two or more display areas, while the less-skilled participants fixated on the ball trajectory and specific players ($F_9 = 6.321$, $p \leq 0.001$, $\eta^2_p = 0.062$), corroborating Roca et al. (2011). Moreover, skilled players generated more condition concepts ($U = 1985.-50$, $z = -4.548$, $p \leq 0.001$, $r = 0.357$) with higher levels of sophistication than their less-skilled counterparts ($U = 1881.00$, $z = -4.423$, $p \leq 0.001$, $r = 0.355$), supporting the works of McPerson and Kernodle (2007) and Botelho et al. (2011). A clear link emerged between visual search behaviors and a more sophisticated knowledge base (expressed through verbal reports of thinking), in the sense that experts attended more to functional spaces, allowing them to capture richer information concerning game problems, which translated into generating more condition concepts of superior sophistication.

Findings highlight the value of using representative task designs to capture performance in situ (Brunswik, 1955; Dicks, Davids & Button, 2009). Although motor responses were not evaluated, they were included in the task, in order to better respect the specific perception-action couplings of real-life situations, therefore capturing more accurately the nature of the skill-based differences (Dicks et al., 2010; Mann et al., 2010).

References

Araújo, R., Afonso, J., & Mesquita, I. (2011) Procedural knowledge, decision-making and game performance analysis in Female Volleyball's attack according to the player's experience and competitive success. *International Journal of Performance Analysis in Sport, 11,* 1-13

Araújo, D., Davids, K., & Passos, P. (2007). Ecological validity, representative design, and correspondence between experimental task constraints and behavioral setting: Comments on Rogers, Kadar, and Costall (2005). *Ecological Psychology, 19*, 69-78.

Behrmann, M., & Ewell, C. (2003). Expertise in tactile pattern recognition. *Psychological Science, 14,* 480-486.

Botelho, S., Afonso, J., Araújo, R., & Mesquita, I. (2011). Relationship between procedural knowledge and player status within the team. Study with female volleyball players during performance of defensive tasks in live-action settings. *Portuguese Journal of Sports Sciences 11*(Suppl. 4), 41-42.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62,* 193-217.

Button, C., Dicks, M., Haines, R., Barker, R., & Davids, K. (2011). Statistical modelling of gaze behaviour as categorical time series: What you should watch to save soccer penalties. *Cognitive Processing, 12,* 235-244.

Dicks, M., Button, C., & Davids, K. (2010). Examination of gaze behaviors under in situ and video simulation task constraints reveals differences in information pickup for perception and action. *Attention, Perception, and Psychophysics, 72,* 706-720.

Dicks, M., Davids, K., & Button, C. (2009). Representative task designs for the study of perception and action in sport. *International Journal of Sport Psychology, 40,* 506-524.

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine, 15,* 988-994.

Ericsson, K. A., & Ward, P. (2007). Capturing the naturally occurring superior performance of experts in the laboratory. *Current Directions in Psychological Science, 16,* 346-350.

Ericsson, K. A., & Williams, A. M. (2007). Capturing naturally occurring superior performance in the laboratory: Translational research on expert performance. *Journal of Experimental Child Psychology: Applied, 13,* 115-123.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Miffin.

Laurent, E., & Ripoll, H. (2009). Extending the rather unnoticed Gibsonian view that 'perception is cognitive': Development of the enactive approach to perceptual-cognitive expertise. In Araújo, D., Ripoll, H., & Raab, M. (Eds.), *Perspectives on cognition and action in sport* (pp. 133-146). New York: Nova Science Publishers.

Mann, D., Abernethy, B., & Farrow, D. (2010). Action specificity increases anticipatory performance and the expert advantage in natural interceptive tasks. *Acta Psychologica, 135*, 17-23.

McPherson, S. (2000). Expert-novice differences in planning strategies during collegiate singles tennis competition. *Journal of Sport and Exercise Psychology, 22*, 39-62.

McPherson, S., & Kernodle, M. (2007). Mapping two new points on the tennis expertise continuum: Tactical skills of adult advanced beginners and entry-level professionals during competition. *Journal of Sports Sciences, 25*, 945-959.

McRobert, A., Ward, P., Eccles, D., & Williams, A. M. (2011). The effect of manipulating context-specific information on perceptual-cognitive processes during a simulated anticipation task. *British Journal of Psychology, 102*, 519-534.

McRobert, A., Williams, A. M., Ward, P., & Eccles, D. (2009). Tracing the process of expertise in a simulated anticipation task. *Ergonomics, 52,* 474-483.

Moreno, M. P., Moreno, A., Ureña, A., Iglesias, D., & Del Villar, F. (2008). Application of mentoring through reflection in female setters of the Spanish National Volleyball team: A case study. *International Journal of Sport Psychology, 39,* 59-76.

North, J., Williams, A. M., Hodges, N., Ward, P., & Ericsson, K. A. (2009). Perceiving patterns in dynamic action sequences: Investigating the processes underpinning stimulus recognition and anticipation skill. *Applied Cognitive Psychology, 23,* 878-894.

Roca, A., Ford, P., McRobert, A., & Williams, A. M. (2011). Identifying the processes underpinning anticipation and decision-making in a dynamic time-constrained-task. *Cognitive Processing, 12,* 301-310.

Williams, A. M., Ford, P., Eccles, D., & Ward, P. (2011). Perceptual-cognitive expertise in sport and its acquisition: Implications for applied cognitive psychology. *Applied Cognitive Psychology, 25*, 432-442.

Williams, A. M., Ward, P., Ward, J., & Smeeton, N. (2008). Domain-specificity, task-specificity, and expert performance. *Research Quarterly for Exercise and Sport, 79,* 428-433.

Judgment and the Probability of Responding

James A. Athanasou
University of Technology, Sydney and Private Practice,
Australia

Contact: athanasou@optusnet.com.au

An earlier historical paper traced the conceptual foundations of the social judgment theory of Brunswik and its transition from measurement through probability of responding to measurement via correlations. The limitations of a correlation approach were considered as inadequate for an idiographic analysis of human responding given that Michell (1990, 2012) has established major problems with measurement in the behavioural sciences. In short, Michell has argued that for the most part there are no attitudinal or Likert-type quantitative units of judgement. Items in typical psychological scales and some judgement studies may lack additivity, which is the essence of true measurement.

An earlier paper suggested a return to the constancy ratio developed by Brunswik as well as to probability of responding. It argued that these have greater validity. The concept of probability is also a feature of two other major behavioural approaches, namely, operant conditioning and Rasch measurement. It is argued that probability has relevance for the development of psychological laws of individual responding and may serve as a unifying construct.

Social judgment based on probability of responding and using the lens model is being explored in the area of (a) rehabilitation judgements where there is a criterion and (b) in the context of religious judgements where there is no obvious criterion.

Probability of rehabilitation judgements where there is a criterion of return to work

This is an ongoing study. To date the Earning Capacity Assessment Form (Shahnasarian, 2010) was administered to 33 accident compensation applicants undertaking a vocational assessment in relation to a legal claim for economic loss. This Earning Capacity Assessment Form contains 14 judgemental criteria. Normally these criteria have been combined to form an overall judgement on a scale form no loss through mild, moderate severe, extremely severe to catastrophic. These criteria are being related to the return to work status of the individual in terms of probability of prediction and in order to determine a set of criteria with the maximum accuracy. Data collection is continuing and it is expected that the study will be completed in December 2012 with a larger sample and a preliminary paper will be available for January 2013. Substantive problems are being experienced with the method of evaluation, the subjectivity involved, the redundancy of the cues and the measurement scales used in the test manual.

Probability of judgement in the context of religious judgements where there is no obvious criterion

This research tests the hypothesis of an ancient writer, St Gregory of Nyssa, that the faith for what is claimed may be derived from the result of other events. The probability of the resurrection of Jesus was claimed to be supported by six events: (a) bringing to life the only son of the widow of Nain; (b) the prediction of the destruction of Jerusalem; (c) Jesus healing Simon's mother-in-law of a high fever and she was able to get up at once and begin to wait upon them; (d) healing the nobleman's son who begged Jesus to come and save his son from death and this was done by his word alone and at a distance and at the exact moment; (e) raising the daughter of Jairus from death; and (f) raising Lazarus from death after four days. In a pilot study, these cues were administered individually to a subject who accepted the probability of the resurrection. This subject thought that each cue provided some support for the idea of the resurrection but ranked them in order of relevance as f, e, d, a, c and b. Pairs of cues were then administered using all possible combinations and a test-retest of only one combination has been held to date. The relative probability of the cues supporting the idea of this religious belief is being investigated. The study is ongoing and it is hypothesised that there will be idiosyncratic components of belief for different individuals who accept the probability or non-probability of a statement of faith.

References

Michell, J. (1990). *An introduction to the logic of psychological measurement.* Hillsdale, NJ: Erlbaum.

Michell, J. (2012). "The constantly recurring argument": Inferring quantity from order. *Theory & Psychology,* 1-17.

Shahnasarian, M. (2010*). Earning Capacity Assessment Form (*2nd edition). Professional manual. Lutz, FL: PAR.

Adaptive Design Optimization for Risky Choice Experiments

Daniel R. Cavagnaro
Mihaylo College of Business and Economics, California State University Fullerton, USA

Richard Gonzalez
Department of Psychology, University of Michigan, USA

Jay I. Myung
Department of Psychology, Ohio State University, USA

Mark A. Pitt
Department of Psychology, Ohio State University, USA

Contact: dcavagnaro@fullerton.edu

Collecting data to discriminate between models of risky choice requires careful selection of stimuli. Models aim to predict choices across a wide range of possible stimuli, but practical limitations force experimenters to select only a handful of them for actual testing. In the traditional paradigm, stimuli are selected by the researcher in advance of the study, often based on the researcher's intuition, and held fixed throughout the experiment. This practice has led some researchers to criticize other decision-making studies for “cherry picking" peculiar examples that lead to particular violations, which could bias those studies against a simple theory that holds very well throughout most of the gamble space. Nevertheless, some stimuli tend to be more diagnostic between two models than do other stimuli, and if the stimulus set does not permit clear differentiation of model predictions then its results may be inconclusive, so the choice of stimuli is critical.

Our recent work (Cavagnaro et al., 2012) considers an algorithmic approach to the selection of decision stimuli that is relatively general in its application. In particular, we provide the theoretical background and a methodological framework for adaptive selection of optimal stimuli for discriminating among models of risky choice. The approach, called Adaptive Design Optimization (ADO), adapts the stimulus in each experimental trial based on the results of the preceding trials. The optimal stimulus at each trial is found by searching the entire feasible stimulus space and identifying stimuli that optimize the discriminability of the models being tested.

This work follows a long tradition of interdisciplinary research on experimental design.

There is a sizable body of work in statistics on formal methods for optimizing the design of an experiment (e.g., Lindley, 1956; Kiefer, 1959; Atkinson & Donev, 1992; Chaloner & Verdinelli, 1995). Drawing on these and other recent developments in statistical computing (Müller et al., 2004; Amzal et al., 2006), Myung and Pitt (2009) developed a design optimization (DO) framework and illustrated its application in discriminating nonlinear models of cognition in two context areas: retention memory

and category learning. The framework was designed as a one-shot process to be performed at the outset of an experiment. Cavagnaro et al. (2010) extended this framework to the case of adaptive design optimization (ADO) in which DO is repeated after collecting only a fraction of all data. Our new research utilizes the ADO framework to optimize decision stimuli for risky choice experiments. We demonstrate its validity with simulations studies aiming to discriminate expected utility, weighted expected utility, original prospect theory, and cumulative prospect theory models.

An abundance of models is one sign of a productive field of inquiry, but this productivity can also be a curse if the models are such close competitors that they cannot be distinguished. To the extent that models can be distinguished, ADO is a new tool that has the ability to overcome such an impasse. It does so by essentially finding vulnerabilities in their data-fitting capabilities and exploiting them until one of the models is shown to be inferior. The adaptive nature of the methodology makes its discrimination process efficient, and although much more development is still needed, our recent demonstrations show that it holds considerable potential.

References

Amzal, B., Bois, F., Parent, E., & Robert, C. (2006). Bayesian-optimal design via interacting particle systems. *Journal of the American Statistical Association, 101*(474), 773-785.

Atkinson, A., & Donev, A. (1992). *Optimum Experimental Designs*. Oxford University Press.

Cavagnaro, D., Gonzalez, R., Myung, J., & Pitt, M. (2012). Optimal decision stimuli for risky choice experiments: An adaptive approach. *Management Science*. Published online before print August 3, 2012, doi: 10.1287/mnsc.1120.1558.

Cavagnaro, D. R., Myung, J. I., Pitt, M. A., & Kujala, J. V. (2010). Adaptive design optimization: A mutual information based approach to model discrimination in cognitive science. *Neural Computation, 22*(4), 887-905.

Chaloner, K., & Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science, 10*(3), 273-304.

Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society. Series B (Methodological), 21*(2), 272-319.

Lindley, D. (1956). On a measure of the information provided by an experiment. *Annals of Mathematical Statistics, 27*(4), 986-1005.

Müller, P., Sanso, B., & De Iorio, M. (2004). Optimal Bayesian design by inhomogeneous markov chain simulation. *Journal of the American Statistical Association, 99*(467), 788-798.

Myung, J. I., & Pitt, M. A. (2009). Optimal experimental design for model discrimination. *Psychological Review, 58*,193-198.

Heuristics and Visual Aids

Mandeep K. Dhami
University of Surrey,
UK

Contact: m.dhami@surrey.ac.uk

There are two areas of research activity that I have been involved in recently which continue my interest in Brunswikian ideas. Both involve applications to the criminal justice and medical domains, and have allowed me to work with some wonderful colleagues. First, beyond further demonstrating the utility of simple heuristics in predicting human judgment, using information search board methods, we have shown that these heuristics can also be potentially valid in describing how doctors search for information before making judgments about prescribing lipid lowering drugs. Using a representatively designed study, we have found that simple heuristics can also be helpful (and adaptive) for experienced burglars when making accurate judgments about residence occupancy (a pre-cursor to burglary).

Second, we have been exploring the effectiveness of visual aids as debiasing techniques for both lay and expert judgment and decision making. To-date, we have found that simple visual displays virtually eliminate framing effects in senior police officers' judgment and decision making, and can help patients from immigrant, non-native language speaking populations to appropriately consider information about proportions when assessing treatment risks. We believe that a potential explanation for the usefulness of visual aids may lie in the fact that the human mind is better evolved to process visual rather than numerical information. Unlike some other debiasing techniques visual aids place the onus on improving cognition beyond the individual decision maker, and so are likely to have a wide impact if implemented.

Below are abstracts for these recent papers:

Snook, B., Dhami, M. K., & Kavanagh, J. (2011). Simply criminal: Predicting burglars' occupancy decisions with a simple heuristic. *Law and Human Behavior, 35,* 316-326.

Rational choice theories of criminal decision making assume that offenders weight and integrate multiple cues when making decisions (i.e., are compensatory). We tested this assumption by comparing how well a compensatory strategy called Franklin's Rule captured burglars' decision policies regarding residence occupancy compared to a non-compensatory strategy (i.e., Matching Heuristic). Forty burglars each decided on the occupancy of 20 randomly selected photographs of residences (for which actual occupancy was known when the photo was taken). Participants also provided open-ended reports on the cues that influenced their decisions in each case, and then rated the importance of eight cues (e.g., deadbolt visible) over all decisions. Burglars predicted occupancy beyond chance levels. The Matching Heuristic was a significantly better predictor of burglars' decisions than Franklin's

Rule, and cue use in the Matching Heuristic better corresponded to the cue ecological validities in the environment than cue use in Franklin's Rule. The most important cue in burglars' models was also the most ecologically valid or predictive of actual occupancy (i.e., vehicle present). The majority of burglars correctly identified the most important cue in their models, and the open-ended technique showed greater correspondence between self-reported and captured cue use than the rating over decision technique. Our findings support a limited rationality perspective to understanding criminal decision making, and have implications for crime prevention.

Dhami, M. K., & Harries, C. (2010). Information search in heuristic decision making. *Applied Cognitive Psychology, 24,* 571-586.

Simple heuristics of the type introduced by Gigerenzer, Todd, and The ABC Research Group (1999) embody principles for information search, stop and decision making. These heuristics suggest that such processes are simple. In an analysis of general practitioners' (GPs) information search and decision-making behaviour when prescribing a lipid lowering drug, we examined whether information search was simple, and whether a heuristic that predicts a simple decision-making process was also accurate at describing information search. We found that GPs' information search behavior was simple in that it demonstrated characteristics of the matching heuristic (e.g. stopping rule). In addition, although the matching heuristic which correctly predicted on average 75% of GPs' decisions used significantly fewer cues on average than the GPs did in the information search task, it was reasonably accurate in describing order of information search. These findings have implications for the validity of simple heuristics describing both information search and decision making.

Garcia-Retamero, R., & Dhami, M. K. (In press). On avoiding framing effects in experienced decision makers. *Quarterly Journal of Experimental Psychology.*

The present study aimed to (1) demonstrate the effect of positive-negative framing on experienced criminal justice decision makers, (2) examine the debiasing effect of visually structured risk messages, and (3) investigate whether risk perceptions mediate the debiasing effect of visual aids on decision making. In two phases, 60 senior police officers estimated the accuracy of a counter-terrorism technique in identifying whether a known terror suspect poses an imminent danger, and decided whether they would recommend the technique to policy makers. Officers also rated their confidence in this recommendation. When information about the effectiveness of the counter-terrorism technique was presented in anumerical format, officers' perceptions of accuracy and recommendation decisions were susceptible to the framing effect: The technique was perceived to be more accurate and was more likely to be recommended when its effectiveness was presented in a positive than negative frame. However, when the information was represented visually using icon arrays, there were no such framing effects. Finally, perceptions of accuracy mediated the debiasing effect of visual aids on recommendation decisions. We offer potential explanations for the debiasing effect of visual aids, and implications for communicating risk to experienced, professional decision makers.

Garcia-Retamero, R., & Dhami, M. K. (2011). Pictures speak louder than numbers: On communicating medical risks to immigrants with non-native language proficiency. *Health Expectations, 14,* (suppl. 1.), 46-57.

Medical risk communication has been infrequently studied in immigrants with limited non-native language proficiency, even though they may be at greatest risk of illness. In a study, we examined to what extent Polish immigrants to the UK have difficulties in understanding treatment risk reduction expressed as ratios either in their native language or in a non-native language (English). We further investigated whether this population can be aided by using visual displays to enhance comprehension. A survey was conducted in the UK involving a sample of Polish immigrants n = 96. Estimates of treatment risk reduction, confidence in estimates, and perceptions of treatment effectiveness were obtained from the sample. When assessing treatment risk reduction, participants often paid too much attention to the number of treated and non-treated patients who died (i.e. numerators) and insufficient attention to the overall number of treated and non-treated patients (i.e. denominators). This denominator neglect was especially noticeable when treatment risk reduction was not expressed in participants' native language. However, provision of visual aids in addition to the numerical information about risk reduction proved to be an effective method for eliminating denominator neglect. The visual aids drew participants' attention to the overall number of treated and non-treated patients and helped them to make more accurate risk estimates. When communicating risks to immigrants with limited non-native language proficiency, we should move beyond the simple, direct translation of health messages that are already being used with the indigenous population to messages that are more appropriate. The use of materials that include visual aids is an effective method of communicating medical risk information to immigrant populations.

Studying Emotional Communication from a Brunswikian Perspective

Chiara Fiorentini, Marcello Mortillaro, Tanja Bänziger, & Klaus R. Scherer
Swiss Centre for Affective Sciences Centre Interfacultaire en Sciences Affectives (CISA),
University of Geneva,
Switzerland

Contact: chiara.fiorentini@unige.ch

In the field of emotion communication it has become increasingly evident that production and perception processes are two sides of a coin and should always be considered jointly. For this reason our group suggested applying a modified version of Brunswik's lens model to the study of emotion communication (Juslin & Laukka 2003; Juslin & Scherer, 2005; Scherer, 2003).

Emotion communication begins when the sender expresses an emotion state by means of several distal cues (distal because they are remote from the observer) likely using multiple expressive modalities (face, voice, gestures, etc.). Distal cues are the result of both cultural/linguistic rules and neurobiological mechanisms that the expresser uses - with or without awareness - to convey his/her state. These distal cues are then perceived by the observer in a relatively modified form as proximal cues (percepts): their degree of similarity to the distal cues depends on various factors, including the quality of the transmission channel and of the sensory system. Crucially, the observer has access only to these proximal cues and based on them he or she makes a probabilistic inference about the emotion that the sender intended to express. The observer uses a range of mechanisms that go from hard-wired pattern of feature detection to appraisals, and schematic decoding, to make this inference (Mortillaro, Mehu, & Scherer, in press)

Our goal is to systematically study the characteristics of distal and proximal cues associated to expressive communications in the different modalities, and their correlations; by using a lens model we want to understand the actual basis of emotion judgments that form an integral part of our everyday life and social interactions.

We recently completed a first set of studies on vocal communication of emotion in which we used the Brunswikian perspective (Bänziger, Patel, & Scherer, 2012a, b). As a first step, we conducted two studies to collect ratings of perceived characteristics of emotional voice and speech. Eight vocal and speech features (loudness, pitch, intonation, sharpness, articulation, roughness, instability, and speech rate) were rated by non-expert (untrained) listeners for two different sets of emotion portrayals. Rater agreement and differences in ratings across emotion categories were examined. Results showed that ratings were consistent across listeners for most scales and also depended on the portrayed emotion. Judgments of four features (loudness, pitch, intonation, and sharpness) were strongly intercorrelated and were also correlated with relevant acoustic descriptors. On the contrary, the ratings of the four other scales (articulation, roughness, instability, and speech rate) were less intercorrelated and were not well predicted by the acoustic

descriptors. Furthermore, despite the use of different portrayals, raters, and procedures, the two studies showed largely convergent results. Overall, these findings therefore suggest that untrained listeners can provide reliable perceptual ratings of at least some aspects emotional speech.

In the second step, we directly assessed the relationship between distal and proximal vocal cues using lens model equations (Bänziger, Patel, & Scherer, 2012b). Acoustic parameters were extracted from vocal emotion portrayals and compared to ratings of voice quality with respect to ecological and functional validity and achievement. The results showed that the vocal descriptors were differentially related to different emotions: the communication of happiness/pleasure was the least well accounted for by the model, whereas the communication of emotional arousal was accounted for almost perfectly both with acoustic descriptors and with ratings of voice features.

These studies clearly point to the usefulness of a Brunswikian approach for the understanding of emotion communication.

The next step will be a similar investigation for emotional facial expressions.

Through the use of specific graphic software we will design synthetic facial expressions based on a precise and quantifiable manipulation of distal cues: in the case of facial expressions distal cues can be thought in terms of individual facial muscle actions. We will collect spontaneous judgments of these expressions: we will ask to naïve observers to verbally describe the perceived appearance of the face as precisely as possible. Rating scales of facial expressions in terms of proximal cues will be developed by using the facial descriptors most frequently used by non-expert judges. Once both sets of cues are collected, the relationship between production and perception processed will be modeled by means of lens model equations.

References

Bänziger, T., Patel, S., & Scherer, K. R. (2012a, in preparation). *Perceived voice and speech characteristics in emotional expression.*

Bänziger, T., Patel S., & Scherer, K. R. (2012b, in preparation) *Understanding the vocal communication of emotion: A Brunswikian approach.*

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

Juslin, P. N., & Scherer, K. R. (2005). *Vocal expression of affect.* In J. A. Harrigan, R. Rosenthal, K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65-135). New York: Oxford University Press.

Mortillaro, M., Mehu, M., & Scherer, K. R. (in press). The evolutionary origin of multimodal synchronization and emotional expression. In E. Altenmüller, S. Schmidt, & E. Zimmerman (Eds.), *Evolution of emotional communication: From sounds in nonhuman mammals to speech and music in man.* New York: Oxford University Press.

Scherer, K. R. (2003). Introduction: Cognitive components of emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldschmidt (Eds). *Handbook of Affective Sciences* (pp. 563 – 571). Oxford University Press, New York and Oxford.

Two Methodological Variants on the Lens Model

Robert M. Hamm
Department of Family and Preventive Medicine,
University of Oklahoma Health Sciences Center,
Oklahoma City, Oklahoma,
USA

Contact: robert-hamm@ouhsc.edu

Question 1: The prototypical Lens Model analysis describes the judgments of one individual. But in studying physicians often it is difficult to get enough cases of the same sort, per physician, and researchers resort to collecting several judgments from each of several physicians. How can such data be analyzed, other than with sheepish apologies?

Question 2: Physicians are often faced with the choice whether to request elaborate (or expensive, or invasive) diagnostic procedures, or to rely on what they know from a clinical exam (that is, talking with the patient and examining the patient's body) when deciding how to manage a disease. How much does the additional information help? Do physicians utilize such information appropriately?

We have applied Lens Model analyses to physician judgment in a way that addresses these two questions. The 2009 Brunswik Society Newsletter http://www.brunswik.org/newsletters/2009news.pdf (page 9) had a report by Robert M. Hamm, Neal V. Dawson, and Rory Ramsey on "Gresham's Law of Judgment: Fancy, Invalid Information Drives out Plain, Valid Information?". We have done more analyses of these data, reported at the annual meetings of the Society for Medical Decision Making in Phoenix, in a poster by Robert M. Hamm, Rory Ramsey, Neal V. Dawson, William A. Whitelaw, Ward W. Flemons, Rollin F. Brant, and Kingman P. Strohl entitled "Physician overutilization of sleep studies in prediction of patient improvement using CPAP" (continuous positive airway pressure) (http://smdm.confex.com/smdm/2012az/webprogram/Paper7002.html).

To address Question 1: The task, as described by Whitelaw et al. (2005) in Am J Respir Crit Care Med, is to predict how well CPAP will help a patient with sleep apnea. We analyzed the predictions made for 262 patients by their sleep specialist physician, from a large body of clinical cues that we reduced to 15 using factor analysis. There were 7 physicians, with varying numbers of patients. Instead of the usual multiple linear regression model, we used ANOVA with "physician" as a 7 level categorical predictor. That handles the physician having different mean predictions. The achievement r_a was low, .36, and ignoring the physician the modeled portion $G*R_e*R_s$ was .21. Accounting for physician means with the ANOVA increased the modeled portion to .24. (The same structure of model was applied to both the ecology - the measure of the patient's actual improvement - and the physician's judgment, of course.) But what about the possibility that physicians differ in the way they utilize the cues? With 15 cues and 7 physicians, it would use up 90 degrees of freedom to include all in the model, and greatly overfit the data. Instead, we constructed all the interaction variables and opportunistically selected the significant

interactions. This capitalizes on chance, and hence may be considered an optimistic estimate of how much can be explained by including physicians' different cue utilization in the model. The model including these physician by cue interactions only increased the model's $G^* R_e^* R_s$ from .24 to .26. This approach gives a specific way for researchers to address the question of how much physicians' judgment differences might impact the Lens Model parameters.

To address Question 2: An earlier study using the Lens Model to address physicians' utilization of a diagnostic procedure on top of the available clinical data was reported by Speroff, Dawson, and colleagues (1989) in Medical Decision Making. They looked at how physicians interpreted clinical data plus non-invasive diagnostic procedures to predict the cardiac risk in 440 patients, compared with the results of the invasive right-heart catheterization (the criterion). They split the modeled portion of the achievement ($G^* R_e^* R_s$ = .29) into two additive portions, one attributable to the clinical data (.26) and the other to the non-invasive tests (.03). A problem with this analysis is that the non-invasive test results are correlated with the clinical data, and so the Lens Model formulation does not clearly measure the diagnostic procedure's additional contribution, or how much the physicians were relying on it.

The sleep study data allow a more satisfactory approach to this issue, because each physician made two predictions for each patient's improvement after the future CPAP treatment: one based on only the clinical data, and the second after an overnight sleep study had been done. We fit these with two Lens Models using the same criterion. The comparison between the two models allows the impact of the new information on judgment accuracy to be seen. (As noted in 2005 and 2009, the r_a went down!). In the analysis of the second judgment, the new predictor was not the overnight sleep study's results. Instead, we used the residual of the prediction of the overnight sleep study from the original clinical variables. That is, we saw how much of the new diagnostic procedure could have been predicted from the clinical data. By removing that, we eliminated the problem that our new variable would be correlated with the other predictors. This allows the weights on the other clinical variables in the ecology model to be exactly the same in the Lens Models for the physician's first and second predictions. Comparison of the weights on these clinical cues in the two judgment models reveals how the physicians change their use of the clinical information (when they had the diagnostic procedure's results, they paid less attention to valid clinical cues). Further, the weight on the new cue reflects only the new information that the diagnostic procedure provides (in the ecology model) and only the use that the physicians made of this new information (in the judgment model). (They paid much more attention to the unique new information than it merited.) Thus this approach handles the ambiguity inherent in the use of a single Lens Model. With two judgments, and two lens models, it is possible to address exactly the questions that are of interest about the value of the diagnostic procedure and the physicians' use of the new and old information.

We anticipate that researchers may find these two methodological variations useful in the analysis of judgments by physicians or experts in other areas.

Different Meanings of the Concept Representativeness

Ken R. Hammond
University of Colorado,
USA

Contact: kenneth.hammond@colorado.edu

Here is the way Kahneman and Tversky introduce the important concept of "representativeness" to their readers:

"This paper explores a heuristic — representativeness — according to which the subjective probability of an event, or a sample, is determined by the degree to which it: (i) is similar in essential characteristics to its parent population; and (ii) reflects the salient features of the process by which it is generated. This heuristic is explicated in a series of empirical examples demonstrating predictable and systematic errors in the evaluation of uncertain events. In particular, since sample size does not represent any property of the population, it is expected to have little or no effect on judgment of likelihood. This prediction is confirmed in studies showing that subjective sampling distributions and posterior probability judgments are determined by the most salient characteristic of the sample (e.g., proportion, mean) without regard to the size of the sample. The present heuristic approach is contrasted with the normative (Bayesian) approach to the analysis of the judgment of uncertainty." (1972, 430-454).

Thus, for Kahneman and Tversky the term "representative" is far different from the statistical term "representative" Brunswik had in mind when he earlier introduced the term in the title of his book "Perception and the Representative Design of Experiments" (Brunswik, 1956). For Brunswik, representativeness is an objective, measurable relation between the cue(s) attached to the objects in the real world (or the parent population) and the cues presented by those objects; there is no reference to a subject in that definition. The purpose of that sentence is to make clear that Brunswik's "representativeness" is independent of the person who is a subject in the experiment. Brunswik's representativeness of the objects in the experiment can be ascertained without reference or interaction with the subject in the experiment; it is present in nature or arranged by the experimenter, and therefore could not be a part of an heuristic approach. (Of course, the subject's perception of the cues is a matter of psychology.)

Thus "representativeness" as Brunswik applied it to the "representative design of experiments" is used in the standard statistical sense, yet it took psychologists roughly 50 years to accept the central idea that if they wanted to generalize their results from the objects in the experiment to any other situation of interest they would have to meet the standard statistical requirement of representativeness, namely, that the objects in the experiment to which the subjects were exposed should be representative of the objects to which the generalization was intended. Regrettably, few psychologists have grasped it yet. In short, it was introduced by Brunswik as an exact parallel to the term representative as it is used to describe a sample of subjects. Therefore, I want the reader to focus on the specific relation between cue

and object, and to note that that there is no place for properties of the subject (person) in that sentence. Note that "representativeness" always was a topic within the framework of the design of the experiment (as the statisticians taught us). Unfortunately, Kahneman and Tversky use it as an heuristic, a psychological concept applying to subjects, without ever explaining the difference in their usage. But Brunswik first used it is simply as an important methodological concept applied to the object side of the experiment, paralleling its use on the subject side; he did it not use it as an heuristic employed by the subject.

This difference is significant for it is precisely that topic of representativeness that is central to allowing generalizations to be drawn by the experimenter, and it is the key to psychologists' failure to investigate generalization rather than accuracy. For if the cue-object relation in the experiment is not representative of the cue-object relation in a specified environment, generalization to that environment is not justified. Indeed, it is precisely at that point that the relation between accuracy and generalization becomes apparent — at least to students of Brunswik.

And that means — to students of Brunswik — that it is the task of the experimenter to select and/or arrange the population of the objects to which the subjects' inferences refer (the parent population) in the experiment, as well as the cues offered to the subject by that sample of objects, and offer a defense of that process. In short, in order to be specific about the test of the accuracy (correspondence) of the subject's inference to the objects of his/her inference, the experimenter must, first, make explicit the properties of the parent population to which the inference is made, and justify the selection of population of objects, and second, justify the process that generates the sample object(s) from which the inference is made to the population of objects. But this topic hardly appears in Kahneman's (2011) recent book; it is focused on the coherence of the decision making process, and notes the several ways in which subjects fail to achieve coherence, and thus fail to achieve a defensible decision in that task. That is why the results are restricted to that task. Thus Kahneman offers demonstrations, rather than scientific experiments.

References

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkley, CA: University of California Press.

Kahneman, D. (2011). *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430-454.

News from the Past Year

Jim Holzworth
University of Connecticut,
USA

Contact: jim.holzworth@uconn.edu

The first publication in multi-year research program along with Tom Stewart and Jeryl Mumpower, concerning how people learn to make selection and detection decisions in the face of uncertainty and different feedback conditions finally appeared in the Journal of Behavioral Decision Making. That paper focused on different feedback conditions (full outcome feedback, partial feedback, and conditional feedback). We are now working on the second paper, focusing on effects of different levels of uncertainty (R_e).

I have also been supervising two doctoral dissertations concerning judgment. In one study, Kris Korbelak examined a joint model of stressor, strain and human judgment. The stressor was task uncertainty (R_e). Study participants (N = 192) were randomly assigned to different levels of task uncertainty (low, medium, high) and completed a series of 150 judgments and five self-reported indications of perceived stress (e.g., Trying to get the job done was a very frustrating experience), coping behavior (e.g., I've been changing my responses to try to do better), frustration with work (e.g., Trying to get the job done was a very frustrating experience), appraisal (e.g., Consider the situation a challenge rather than a problem), and task engagement (e.g., I was immersed in the task). Results indicate that task uncertainty was directly related to perceived stress, coping behavior and judgment (cognitive control and achievement). Coping behavior did not moderate the effect of perceived stress on judgment.

The second study, being conducted by Scott Ryan, involves a collection of experiments concerned with forecasting. Experiments are designed to test effects of knowledge and individual difference measures on judgment accuracy. It is hypothesized that the most accurate judgments emerge when individuals are selected based on knowledge, confidence, and skill on other forecasting tasks. It is also hypothesized that individuals will be most accurate when provided relevant information cues. Two earlier (pilot) studies indicated that participants made accurate judgments when their judgments were aggregated, but only if they were provided relevant information cues. When individuals were provided both relevant and irrelevant cues, they used the cues inappropriately and made inaccurate judgments. Individuals were not able to distinguish between relevant and irrelevant cues, but instead used all cues provided. Results also showed that for difficult judgments, groups that were selected based on confidence outperformed groups not selected on confidence. These studies lead to more specific hypotheses. Individuals appear to be accurate with respect to weighing relevant cues, but not accurate with respect to selecting which cues are relevant. In order to help individuals properly select cues, a third experiment will provide individuals with multiple sets of cues and outcomes in order to test whether individuals can use this information to learn which cues are

relevant. Individuals will also be asked for their knowledge, confidence, and other individual difference measures. It is hypothesized that knowledge, success on other predictions, and confidence will all be predictive of accuracy.

Judgment Gap Analysis between Service Provider and Consumer for Service Design

---

Seung-Kweon Hong
Korea National University of Transportation,
Korea

Contact: skhong@ut.ac.kr

It is important to understand the good service that service providers and customers are thinking. If there is judgment gap between service providers and customers, it would cause an unsatisfactory service. The judgment gap should be thoroughly investigated for a good service design. The aim of this paper (Hong, 2012) was to introduce a method that can measure and analyze the judgment gaps between service providers and customers. Lens model (Tucker, 1964) was applied to analyze judgment gaps between service providers and customers. As a case study, a library lending service was selected. 5 librarians and 15 customers participated in the experiment that investigates their judgments on a good service. Five dimensions (e.g. Reliability, Assurance, Tangibles, Empathy, Responsiveness) that were proposed to measure customers' expectation and perception (Parasuraman, et al., 1988) were provided to participants as decision cues. Cue weighting policies of consumers and service providers were similar, except that consumers gave higher weight on tangibility than service providers. Service providers and consumers had a good knowledge on the service quality, but they could not well apply the knowledge to judge it. The lens model may be used to analyze judgment gaps between service providers and consumers in the other service areas. The decision cues that were used in this study can be changed, depending on the characteristics of the target service.

References

Hong, S.-K. (2012). Judgment gap analysis between service provider and consumer for service design. *Journal of the Ergonomics Society of Korea, 31*(1), 77-83.

Tucker, L. R. (1964). A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch and by Hammond, Hursch, and Todd. *Psychological Review, 71*, 528-530.

Parasuraman, A., Zeithaml, V., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64,* 12-40.

A Model Study on Emotional Communication in a Mono-Brand Fashion Store – Application of the Lens Model in the Fashion Industry

Zhiqing Jiang & Shin'ya Nagasawa
Graduate School of Commerce, Waseda University,
Tokyo, Japan

Contact: Zhiqing-jiang@ruri.waseda.jp, nagasawa@waseda.jp

The purpose of this paper is to propose a model that analyses the visual efficiency of emotional communication from brand manager to consumers in a mono-brand fashion store. Before setting up the model, an investigation of relevant literatures in marketing and psychology was done in order to put forward a hypothesis of the model. Based on recommendations of the extant research, the scale used to measure emotional communication in this study mainly relied on measuring brand personality. Then, an empirical, statistical study was conducted of Tod's in China for verifying the feasibility of the emotional-communication model. The results show that the emotional-communication model works to explain the visual efficiency of emotional communication in a mono-brand fashion store; a set of regression equations of estimated expectation is also provided on both brand manager and consumers' part. This research has presented a new perspective on fashion brands and on managers, consumers, and brand stores as a single mechanism.

INTRODUCTION

In recent years, one important symbolic brand association in marketing research has been brand personality, which is defined formally as "the set of human characteristics associated with a brand", and human perceptions are the basis of individual behavior, attitudes and beliefs, as well as physical and demographic characteristics (Asker, 1997). Unlike "product-related attributes" serving a utilitarian function for consumers, brand personality tends to serve a symbolic or self-expressive function (Keller, 1993), which contributes to the emotional effects of the brand. Emotion arguably provides the means to coordinate the diverse mental and physical components required to respond to the world in a coherent fashion (Cosmides & Toody, 2000). Meanwhile, on the part of consumers, it is the store rather than the brand that acts as the activation area. Different groups of consumers prefer different types of retail stores (Finn & Louviere 1990). In the fashion industry, this phenomenon is extremely significant because the industry has undergone a slow, general shift in the last decade away from production towards retailing and services. This has been one of the most important aspects of the fashion sector's evolution (Tartaglione, 2005). The communication of emotion in a mono-brand fashion store has been chosen as the subject of this thesis because the perceived emotional essence of a brand influences the degree of consumer brand loyalty; in other words, the emotional experience can contribute to the brand's development, while, conversely, brand personality can offer consumers various emotional

experiences. Therefore, how consumers feel about a brand in a mono-brand fashion store is worth discussion.

References

Asker, J. L. (1997). Dimensions of brand personality. *Journal of Marketing Research, 35,* 347-356.

Cosmides, L., & Toody, J. (2000). *Consider the source: The evaluation of adaptations for coupling and metarepresentation*. In D. Sperber (Ed.), *Metarepresentations* (pp. 53-115). New York: Oxford University Press.

Finn, A., & Louviere, J. (1990). Shopping-center patronage models fashioning a consideration set and segmentation solution. *Journal of Business Research, 21*, 259-275.

Keller, K. L. (1993). Conceptualizing, measuring and managing customer-based brand equity. *Journal of Marketing, 57,* 1-22.

Tartaglione, C. (2005). *Il sistema Moda verso il rinnovamento strategico in un mercato dove sono strutturalmente cambiate le leve della competizione*. Mimeo, Rome, Italy.

Meta-Analysis Research under the Critical Lens: A Missing Brunswik Link?

Esther Kaufmann
University of Zurich / University of Fribourg
Switzerland

Contact: esther.kaufmann@gmx.ch

During the last 30 years, ever since meta-analysis has been used for evaluating research results, a number of variants has been introduced such as the Hedges–Olkin (1885), Rosenthal-Rubin (see Rosenthal, 1991), and the Hunter-Schmidt (2004) method. These meta-analysis approaches differ in varying respects, such as for example, estimated effect size, the applied model being fixed or random, the use of correlation procedures, how to identify eventual moderator variables or if they recommend applying any publication bias estimation.

In the meantime, the different meta-analysis approaches are targets for critiques or evaluation processes. For example Ioannidis (2010) highlighted meta-analysis research as the art of getting it wrong and concluded:

> Meta-analysis, as a prototype research design is a fine tool for making mistakes and sometimes for recognizing the biases that underlie them. …I would welcome more empirical work on the performance of various tests and approaches that we have in detecting biases and true effects since I think that their performance is often suboptimal and over-valued (p. 180).

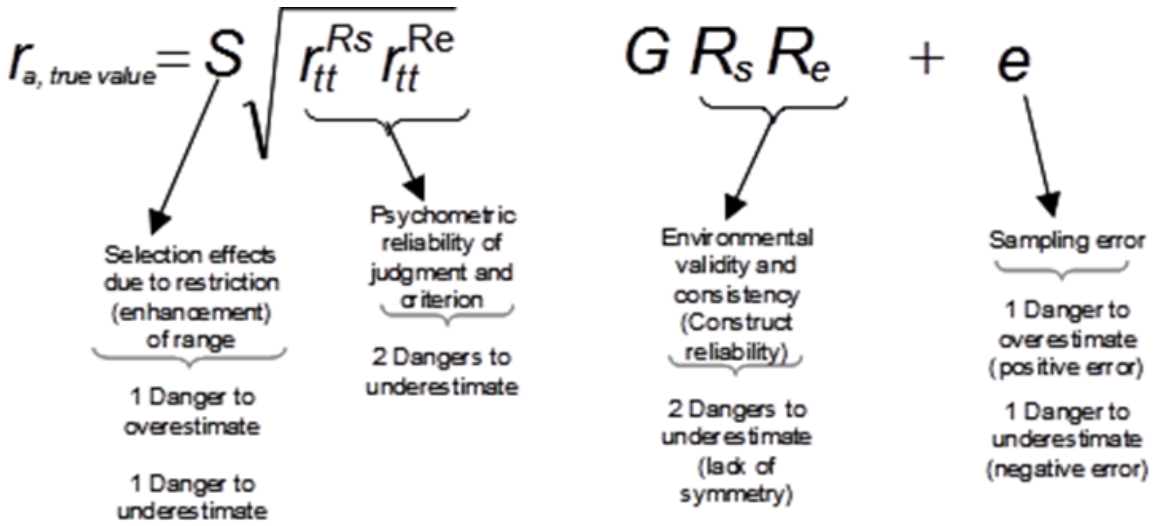
As a response to this we have undertaken a “Brunswikian evaluation” of numerous meta-analysis approaches.

Different meta-analysis approaches have been compared and evaluated in a number of studies. For example, Field (2001), conducted two Monte Carlo studies to compare three meta-analytic approaches. This study shows that the Hunter-Schmidt method ‘tends to provide the most accurate estimates of the mean population effect size when effect sizes are heterogeneous, which is the most common case in meta-analytic practice’ (Field, 2001). Besides these simulation approaches, studies on real

data also support the use of the random-effect model applied by the Hunter-Schmidt method (Kisamore & Brannick, 2008).

Since Wittmann (1988, 2009) linked the Tucker's lens model equation to the psychometric Hunter-Schmidt approach, the Brunswikian symmetry concepts have been introduced in the development of meta-analysis designs. Unlike other meta-analytical approaches the psychometric Hunter-Schmidt variant applies up to 11 artifact corrections to try to estimate true values without any possible artifacts. Most studies on meta-analysis evaluation using the Hunter-Schmidt approach have not linked it to the Brunswikian symmetry concepts, i.e. they have not applied the complete psychometric variant of this approach. For example, the comparative study by Aguinis, Sturman and Pierce (2008), who modeled the three most influential sources of artifact variance, sampling error, measurement error and range restriction, showed that the Hunter-Schmidt method was more precise than the Hedges-Olkin approach when it came to point estimates and homogeneity tests to prevent Type 1 error rates.

With the following equation (Wittmann, 1988, 2009), applied in our study (Kaufmann, 2010), we want to demonstrate the consequences if the Hunter-Schmidt approach is incomplete, i.e., when the psychometric Brunswikian link is neglected. The equation below also demonstrates that if no artifact corrections are included in an evaluation of meta-analysis studies, there is a tendency to underestimate- or overestimate the estimated values even when the most important artifacts are corrected as, for example, in the study by Aguinis, Sturmann and Pierce (2008). There is still a tendency to over- or underestimate the "true" values, due to neglect of the Brunswik symmetry concept (Wittmann, 1988, 2009).



Consequently, we conclude according to Wittmann's equation that there is a two to six (2 to 6) tendency to underestimate the results of meta-analysis evaluations if they do not include a complete psychometric Hunter-Schmidt approach or if this Brunswik link is neglected.

Hence, research on meta-analysis should be aware of this Brunswik link – the complete psychometric Hunter-Schmidt meta-analysis approach and include it in their evaluation work – to accurately estimate the value of the different meta-analysis approaches and to apply them fruitfully. Finally, our contribution is an attempt to

show that Brunswikian concepts could be applied to the evaluation of meta-analysis approaches thereby serving as a response to Ioannidis' (2010) critique.

References

Aguinis, H., Sturman, M. C., & Pierce, C. A. (2008). Comparison of three meta-analytic procedures for estimating moderating effects of categorical variables. *Organizational Research Methods, 11*(1), 9-34.

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods, 6*(2), 161-180.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Ioannidis, J. P. A. (2010). Meta-research: The art of getting it wrong. *Research Synthesis Methods, 1*, 169-184.

Kaufmann, E. (2010). *Flesh on the bones: A critical meta-analytical perspective of achievement lens studies*. (Doctoral dissertation, MADOC: University of Mannheim). Retrieved from http://madoc.bib.uni-mannheim.de/madoc/volltexte/2010/2892/

Kisamore, J. L., & Brannick, M. T. (2008). An illustration of the consequences of meta-analysis model choice. *Organizational Research Methods, 11*(1), 35-53.

Rosenthal, R. (1979). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

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

Wittmann, W. W. (2009). Evaluationsmodelle. In H. Holling (Ed.). *Enzyklopädie der Psychologie. Themenbereich B Methodologie und Methoden. Serie IV Evaluation - Band 1. Grundlagen und statistische Methoden der Evaluationsforschung* (pp. 59-98). Göttingen: Hogrefe.

The External Validity of Research in the Psychological Laboratory

---

Gregory Mitchell
University of Virginia,
USA

Contact: greg_mitchell@virginia.edu

A little over ten years ago, Anderson, Lindsay and Bushman (1999) collected meta-analyses in which effects in the laboratory were compared to effects in the field (e.g., does goal difficulty have the same effect on task performance in the field as in the lab?) and computed the correlation between 38 pairs of meta-analytic effect size estimates for laboratory versus field settings. After finding a large correlation (r = .73) between lab and field effects, Anderson and colleagues concluded that results from psychological laboratories were externally valid. I recently undertook a replication and extension of Anderson et al. (1999) based on 82 meta-analyses reporting 217 pairs of effect size estimates from laboratory and field studies of a wide range of phenomena. With this expanded database, I was able to examine external validity by psychological subfield and topic of study, and my results were more qualified than those of Anderson et al. (1999). I replicated their overall result by finding a large correlation between effects from the laboratory and those from the field (r = .71 when an outlier is excluded), but this result depended on the extremely high correlation of

laboratory and field effects from I-O psychology (with I-O effects excluded, the overall correlation drops to r = .55). External validity differed across psychological subfields and across research topics within each subfield, and all subfields showed considerable variation in the relative size of effects found in the laboratory versus the field (i.e., even though the effects often moved in parallel, the magnitudes of the effect found in the lab and the field often differed greatly). External validity also differed by effect size: Small laboratory effects were less likely to replicate in the field than larger effects. Perhaps most troubling, effects from social psychology laboratories changed signs in the field over 20% of the time, compared to much smaller percentages for other subfields.

Although my study cannot pinpoint the sources of variance between I-O psychology and those subfields producing less externally valid results, one likely source is a difference in guiding design principles. I-O psychologists who use laboratory studies tend to be quite concerned with external validity and persuading those in the field that they have faithfully simulated the settings and tasks under study. Social psychologists and those in other subfields are often guided by the "psychological realism" design principle (Aronson, Wilson, & Akert, 1994, p. 58), which seeks to engage psychological processes from everyday life but often with novel or unrepresentative stimuli. Thus, we should not be too surprised when different configurations or types of stimuli outside the laboratory lead to different effects, or even opposing effects. The irony is that the very subdiscipline that supposedly studies the power of "the situation" often utilizes an experimental design approach that treats environmental variables as if they were interchangeable.

Available online at:
http://pps.sagepub.com/content/7/2/109

References

Anderson, C. A., Lindsay, J. J., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science, 8,* 3–9.

Aronson, E., Wilson, T. D., & Akert, R. M. (1994). *Social psychology: The heart and mind*. New York, NY: Harper Collins.

News from the Past Year

Jeryl L. Mumpower
U.S. National Science Foundation, on leave from Texas A&M University, USA

Contact: jmumpower@tamu.edu, jmumpowe@nsf.gov

The first paper in our multi-year research program in which Tom Stewart, Jim Holzworth and I are attempting to learn more about how people make selection and detection decisions (e.g., to decide whether to hire someone or whether the patient has a disease) in the face of uncertainty and different feedback conditions finally appeared in print. The paper entitled "Learning to Make Selection and Detections Decisions: The Role of Base Rate and Feedback" was just published by the Journal of Behavioral Decision Making. We are now working on the next paper in our research program, which will focus on the effects of uncertainty on performance in such selection and detection decisions.

Other papers that I am working on include (1) a paper with Arnie Vedlitz and others at Texas A&M University entitled "Psychometric and Demographic Predictors of the Perceived Risk of Terrorist Threats and Willingness-to-Pay for Risk Management Programs;" (2) a paper with Mike Lindell and others at Texas A&M University entitled "Perceptions and Expected Responses to a Water Contamination Emergency"; (3) a paper with Gary McClelland at the University of Colorado, Boulder, entitled "A Signal Detection Analysis of Racial and Ethnic Disproportionality in the Referral and Substantiation Components of the Child Welfare Services System."

Finally, I'm unsure how much of this contemplated research I'll actually get done next year because in September 2012 I accepted a two-year IPA (Intergovernmental Personnel Act) assignment to serve as Division Director for Social and Economic Sciences at the U.S. National Science Foundation.

Applications and Extensions of Lens Model Analyses in Personality Psychology

Steffen Nestler & Mitja D. Back.
University Münster,
Germany

Contact: steffen.nestler@wwu.de; mitja.back@wwu.de

Here, we want to give a brief overview of recent applications and extensions of Brunswik's lens model (e.g., Brunswik, 1956) in the domain of personality psychology (see also Nestler & Back, 2012 for a more thorough description). One of the most interesting phenomena in personality psychology is that individuals (called perceivers) are able to correctly judge the personality traits of other individuals (called targets) even if they have not interacted with them (see Funder, 2012). Perceivers, for example, made valid trait ascriptions when they were exposed to brief video clips (Borkenau & Liebler, 1992), Online Social Network profiles (Back, Stopfer, et al., 2010), have written short stories (Küfner, Back, Nestler, & Egloff, 2010), or the e-mail addresses of the targets (Back, Schmukle, & Egloff, 2008). To explain this accuracy, a number of personality psychologists (e.g., Borkenau & Liebler, 1992; Funder, 1999; Gosling et al., 2002; Naumann et al., 2009; Nestler et al., 2012) invoke the lens model. The basic idea is that perceivers use a set of observable attributes (e.g., the stylishness of a person) to judge the personality of the targets (e.g., extraversion). If these attributes are related to the trait (cue validity), and perceivers use them for their judgments (cue utilization), then accuracy will result. Briefly, thus, people make accurate personality inferences when they are sensitive to cues that are valid for the trait in question.

In our own research, we also use the lens model to examine the accuracy of people's personality inferences. In addition, we apply it to understand the accuracy of people's metaperceptions (e.g., how well X knows that Y thinks that X is extraverted) and the effects of personality on liking (e.g., why X likes Y). Regarding metaperceptions, for example, we speculated and found evidence that a metaperception of a target is accurate to the extent that we use the same observable behaviors to infer the judgment of how others view them as the perceivers. Regarding liking, a trait leads to popularity (being liked by others) if it expresses in behavioral cues that others perceive and judge as positive (Back, Schmukle, & Egloff, 2010, 2011; Küfner, Nestler, & Back, in press).

In addition to these extensions of the lens model to other applications, we also differentiated the lens by integrating it with current dual process models. The basis of this differentiation was that all parts of the lens model (the actual trait of the person to be judged, the observable features, the criterion; the judgment of the perceiver; i.e., the judgment; the observable behaviors, i.e., the cues) can be conceptualized on a controllable (roughly, explicit) and an uncontrollable (roughly, implicit) level. Regarding the cues, to begin with, behaviors that are related to a trait can be deliberately generated (e.g., the content of what people say) or they occur more spontaneously (e.g., the loudness of the voice). The trait of a target, second, can be

assessed by explicitly asking the person in a personality questionnaire or by asking the target to complete indirect tests of personality such as the implicit association test. Finally, perceivers can be asked to make an intuitive, spontaneously generated personality judgment or to deliberately think about it. We think that this differentiated lens model, we called it the dual lens model (see Hirschmüller, Egloff, Nestler, & Back, in press), allows us to better understand the accuracy of personality judgments. We also used it to examine a number of exciting research questions such as whether persons are more accurate in judging the targets explicit self-views compared to their implicit ones.

Besides this suggested differentiation, we also integrated the assumptions of the lens model with newer accounts of the hindsight bias that treats it as a by-product of knowledge updating. Specifically, in our integrated model (see Nestler, Egloff, Küfner, & Back, 2012) we posit that people exhibit a hindsight bias as they (a) learn to more accurately perceive the cue values of the to be judged objects, (b) are more sensitive to valid cues, and/or (c) learn to use their cue knowledge more consistently. Empirical findings showed that both the utilization of more valid cues and changes in cue perceptions—but not changes in the consistency with which cue knowledge is applied—were associated with hindsight effects. One very interesting implication of this model is that it can in principal explain where knowledge of cues people use stems from, and—more generally—that it suggests a way to improve the accuracy of trait ascriptions. Specifically, the assumptions of the model entail that when people exhibit a hindsight bias they should judge the personality of new targets more accurately.

To sum up, in personality psychology the lens model is typically used to explain the accuracy of personality judgments. We believe that it can also be fruitfully employed to understand the accuracy of people's metaperceptions and their liking judgments. Also, the lens model can be differentiated by combining it with current research on dual process models. Finally, for an even more comprehensive understanding of interpersonal judgments, the lens model can be integrated with research in hindsight bias. In a nutshell, thus, Brunswik's lens model is an invaluable conceptual framework and a very flexible tool for personality psychologists.

References

Back, M. D., Schmukle, S. C., & Egloff, B. (2011). A closer look at first sight: Social relations lens model analyses of personality and interpersonal attraction at zero acquaintance. *European Journal of Personality, 25*, 225-238.

Back, M. D., Schmukle, S. C., & Egloff, B. (2010). Why are narcissists so charming at first sight? Decoding the narcissism-popularity link at zero acquaintance. *Journal of Personality and Social Psychology, 98*, 132-145.

Back, M. D., Schmukle, S. C., & Egloff, B. (2008). How extraverted is honey.bunny77@hotmail.de? Inferring personality traits from email addresses. *Journal of Research in Personality, 42*, 1116-1122.

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science, 21,* 372-374.

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62*, 645–657.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, CA: University of California Press.

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.

Funder, D. C. (2012). Accurate personality judgments. *Current Directions in Psychological Science, 21*, 177-182.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*, 379-398.

Hirschmüller, S., Egloff, B., Nestler, S., & Back, M. D. (in press). The dual lens model: A comprehensive framework for understanding self-other agreement of personality judgments at zero acquaintance. *Journal of Personality and Social Psychology.*

Küfner, A. C. P., Back, M. D., Nestler, S., & Egloff, B. (2010). Tell me a story and I will tell you who you are! Lens model analyses of personality and creative writing. *Journal of Research in Personality, 44,* 427-435.

Küfner, A. C. P., Nestler, S., & Back M. D. (in press). The two pathways to being an (un-)popular narcissist. *Journal of Personality*.

Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin, 35*, 1661-1671.

Nestler, S., & Back, M. D. (under review). Lens Modeling: Understanding interpersonal judgments at zero acquaintance. *Current Directions in Psychological Science.*

Nestler, S., Egloff, B., Küfner, A. C. P., & Back, M. D. (2012). An integrative lens model approach to bias and accuracy in human inferences: Hindsight effects and knowledge updating in personality judgments. *Journal of Personality and Social Psychology, 103,* 698-717.

How do People Judge Risks: Availability Heuristic, Affect Heuristic, or Both?

Thorsten Pachur & Ralph Hertwig
School of Social Sciences, Max Planck Institute for Human Development, Berlin, Germany

Florian Steinmann
University of Basel, Switzerland

Contact: pachur@mpib-berlin.mpg.de

How do people assess the risk of potentially harmful events in the environment, such as cancer, hurricanes, terrorist attacks, or car accidents? Two prominent proposals of mental mechanisms underlying risk judgments are the availability heuristic (Tversky & Kahneman, 1973) and the affect heuristic (Finucane, Alhakami, Slovic, & Johnson, 2000). Despite their prominence, however, these two heuristics have never been tested against each other. Moreover, how people judge the risk of an event can be measured in several ways — such as perceived frequency of occurrence, value-of-a-statistical-life, or subjective risk; depending on which measurement is used, different mechanisms might be triggered. We conducted a comparison between different implementations of the availability heuristic and the affect heuristic to see which accounts for judged risk better and how ecologically valid each is. To that end, we used a representative design and presented participants with large and varied sets of health risks (including rare and frequent ones). The availability heuristic and the affect heuristic were operationalized based on precisely defined definitions of the availability heuristic and the affect heuristic — thus addressing the often-criticized lack of precision in the original descriptions of these models.

In Study 1, participants indicated for pairs of risks which risk occurs more frequently (frequency judgments); in addition, we assessed for which risk in each pair

people thought that a higher amount of public money should be invested to prevent a fatality due to the risk (value-of-a-statistical-life; VSL). As previous research indicated an association between affect and individualized representations of risk (e.g., Slovic, Monahan, & MacGregor, 2000), we reasoned that the role of affect might be more pronounced in the VLS judgments. As a measure of availability, each participant indicated the number of people in their social circles who had died of each risk. Based on this measure, we tested availability-by-recall (Hertwig, Pachur, & Kurzenhäuser, 2005), a version of the availability heuristic which predicts that an event is judged to be more risky if one can recall a higher number of people. As measure of affect, we determined for each risk its dread score based on the risk questionnaire developed in the context of Slovic and colleagues' psychometric paradigm (e.g., Fischhoff, Slovic, Lichtenstein, Read, & Combs, 1978). Our implementation of the affect heuristic predicts that an event is judged to be more risky if it has a higher dread score.

Figure 1 shows the percentage of correctly predicted judgments of availability-by-recall and the affect heuristic, separately for the frequency judgments and the VSL judgment. The results suggest that the role of availability and affect differs between the two types of risk judgments. For frequency judgments, availability-by-recall clearly outperformed the affect heuristic whereas for VSL judgments the two mechanisms tied. In general, the predictive power of the affect heuristic was higher in VSL judgments than in frequency judgments whereas for availability-by-recall the pattern was reversed. Moreover, based on the actual mortality statistics of the different risks, we were able to determine the ecological validity of availability-by-recall and the affect heuristic for frequency judgments. With 71.9%, availability-by-recall pointed to the actually more frequent risk more often than the affect heuristic (47.6%).

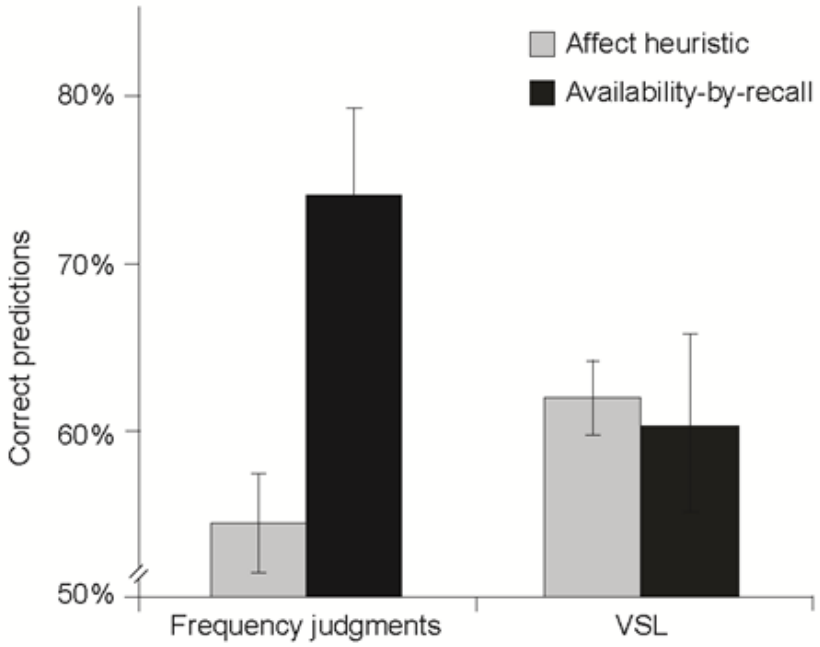


Figure 1. The performance of availability-by-recall and the affect heuristic in predicting people's risk judgments—either in terms of perceived frequency or in value-of-a-statistical life (VSL).

Study 2 replicated these results using a different set of risks and including an alternative version of the availability heuristic. Specifically, we also considered an implementation that included the number of instances of each risk a person could recall from the media — thus augmenting the definition of availability to include instances from both people's personal and their virtual social circles. A further extension was that in addition to frequency and VLS judgments, we also asked a group of people to judge the risks in terms of their subjective risk (i.e., without defining "risk" any further). Overall, the results again indicated a stronger role of availability in frequency judgments than in VLS and subjective risk judgments; in the latter two, in turn, the role of affect was increased (relative to frequency judgments).

In addition to pitting, to our knowledge for the first time, the arguably most prominent models of risk judgments against each other and elucidating their differential roles in different measures of judged risk, our results have several practical implications. First, conclusions concerning people's calibration to the risk ecology depend on how risk perception is measured. As the number of recalled instances in one's social network is an ecologically valid cue for inferring objective mortality, asking people to judge risks in terms of their frequencies is likely to lead to better calibration than asking people to indicate their subjective risk or VSL judgments (which seems to trigger a less ecologically valid cue, namely dread). Second, our results suggest an important and neglected dimension in the design of public health campaigns. Rather than providing people with aggregate frequency statistics about risks, an alternative and potentially more effective way to educate the public about a specific risk could be to encourage people to trust their own experience concerning the risk garnered from their proximate social network. As our results show, if a risk has claimed fewer victims in one's own social circle than another risk, then — on average — this asymmetry is a rough but good proxy for the risks' relative population frequencies.

References

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 13,* 1-17.

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences, 9*, 127-152. doi: 10.1007/BF00143739

Hertwig, R., Pachur, T., & Kurzenhäuser, S. (2005). Judgments of risk frequencies: Tests of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 621- 642.

Pachur, T., Hertwig, R., & Steinmann, F. (2012). How do people judge risks: Availability heuristic, affect heuristic, or both? *Journal of Experimental Psychology: Applied, 18,* 314-330.

Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior, 24,* 271-296. doi:10.1023/A: 1005595519944

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Prediction Why Internal Validity is not Prior to External Validity: Paper presented at PSA 2012

Johannes Persson & Annika Wallin
Lund University,
Sweden

Contact: johannes.persson@fil.lu.se

Donald T. Campbell introduced the concepts internal and external validity in the 1950s. Without a doubt the concepts captures two features of research scientists are aware of in their daily practice. Researchers aim to make correct inferences both about that which is actually studied (internal validity), for instance in an experiment, and about what the results 'generalize to' (external validity). In addition, it is often claimed that one of the two is prior to the other. And the sense in which internal validity is often claimed to be prior to external validity is both temporal and epistemic, at least. The claim about temporal priority is that we first make inferences about the local environment under study before making inferences about the surrounding world. The claim about epistemic priority is that we come to know the local environment before we come to know the surrounding world.

In Persson and Wallin (2012) we problematize the relation between external and internal validity. Our claim is that the two types of validity are deeply intertwined. We argue in favour of the part of the claim that is in conflict with the idea behind the internal/external distinction. The argument is directed at showing that internal validity understood as prior to external validity has, at least, three epistemologically problematic aspects: experimental artefacts, the implications of causal relations, and how the mechanism is measured.

First, the possibility of experimental artefacts in itself questions whether there can be such a thing as an internally valid inference. That clearly depends on whether the methods we use guarantee that we see clearly, i.e. that what we see in the local environment is not in fact an artefact of something else. But some well-known "internally valid" results have in fact been produced by, for instance, the method of randomization or measurement used.

Second, whether there can be an internally valid inference also depends on the nature of what is inferred to. Normally, the inference is considered to be causal (but see, Campbell, 1986, p. 76). Now, there are many concepts of causation. Some of these are clearly of a kind that does not support inferences that are primarily internal. For instance, someone operating with a notion of causation similar to one of those that Kant, Hume, or Mill relied on should judge internally valid inferences to causal matters impossible. For each of those causal concepts the implications of causation, regardless of whether it has to do with the notion of sufficiency or necessity, go beyond the local environment. If there is a causal relation in the local environment it follows that this holds also outside this environment. And, trivially, it holds that if it does not hold outside the environment it cannot hold inside either. Hence such concepts of causation warrant neither the alleged temporal nor epistemological priority of internal validity.

Thirdly, how mechanisms are measured has a strong impact on the results obtained. This might be argued to be a special case of experimental artefacts, but it has particularly interesting implications for the internal-external validity distinction. For instance, Poulton (1975) presents a number of different range effects demonstrating how the order in which stimuli is presented affects the result, or the type of mechanism that is being observed (an "unbiased" perceptual judgment, or judgments mediated by range effects – in themselves mechanisms). Since participants' pre-conceptions of what the range of stimuli is will affect their responses, the "external validity" of the stimuli (in this context how well the range it introduces, or the range the experimenter assumes, matches participants' pre-conceived range of stimuli) determines whether the results obtained in the laboratory correctly capture the features of the mechanism operating there. Hence, in cases like these, external validity is a requirement for internal validity. Note that this potentially false estimate of the function has perfect internal validity. Given the range, the stimuli really do cause the response, and we have a fair grasp of what the mechanisms are.

Egon Brunswik's insistence that psychology has to focus on organism-environment relationships (Brunswik, 1957), underscores the complicated relationship between external and internal validity. If the aim of an experiment in psychology is to understand the functioning of different psychological mechanisms (in the form of stimulus-response relations), then the quality of this finding is just as dependent on whether the psychological mechanism has been properly activated as it is on whether the results can be replicated. This is not only a question about how the result will generalize to other settings (external validity) – it is a question about whether a proper result has at all been generated (internal validity). Thus, for psychological mechanisms that can be assumed to have an adaptive character, external validity (or certain aspects of it) appears to be prior to internal validity: It is more important that an experiment measures what it aims to measure than that the result is internally valid.

References

Brunswik, E. (1957). Scope and aspects of the cognitive problem. In J. S. Bruner, E. Brunswik, L. Festinger, F. Heider, K. F. Muenzinger, C. E. Osgood, & D. Rappaport (Eds.), *Contemporary Approaches to Cognition* (pp. 5-31). Cambridge: Harvard University Press.

Campbell, D. T. (1986). Relabeling internal and external validity for applied social sciences. *Advances in Quasi-Experimental Design and Analysis. New Directions for Program Evaluation*, *31,* 67-77.

Persson, J., & Wallin, A. (2012). *Why internal validity is not prior to external validity.* In: [2012] Philosophy of Science Assoc. 23rd Biennial Mtg (San Diego, CA), PSA 2012 Contributed Papers. Phil Sci Archive. Retrieved from http://philsci-archive.pitt.edu/9171/

Poulton, E. C. (1975). Range effects in experiments on people. *The American Journal of Psychology, 88*(1), 3–32.

Dual-Processes in Learning and Judgment: Evidence from the Multiple Cue Probability Learning Paradigm

Jonathan J. Rolison
School of Psychology,
Queen's University Belfast,
UK

Contact: j.rolison@qub.ac.uk

Kenneth Hammond introduced to the field of human judgment a distinction between intuitive and analytic types of thinking (Hammond, 1996). Modes or styles of cognition, according to Hammond, could be placed on a continuum with analytic and intuitive styles polarized at each end. Characterized as a blend of intuitive and analytic thinking styles, Hammond's view of judgment highlights the various manners in which expertise is acquired and expert knowledge used. For instance, consider how a combination of professional training and intuitive gut-feeling could raise the suspicions of an airport customs officer that a passenger may be carrying drugs. While Hammond was cautious not to dichotomize intuitive and deliberative thinking, at a broad theoretical level, some authors argue for separate implicit and explicit learning systems that underlie intuitive and analytic thought.

My colleagues and I explored the degree to which implicit intuitive and explicit analytic processes are involved in learning and judgment in a multiple cue environment (Rolison, Evans, Dennis, & Walsh, 2012). Using the multiple cue probability learning paradigm, we trained participants on a set of novel cues for predicting a criterion whilst monitoring their explicit beliefs for each of the cues on a trial-by-trial basis during training. Each cue could predict the criterion either positively, negatively, or be irrelevant to it. Participants' explicit ratings of the cues recording during training indicated that participants were engaged in analytic reasoning about the cues, and explicitly tested hypotheses against the cue information. However, when we examined their actual judgments we found that participants' beliefs about negative cues but not their beliefs about positive cues fed into their actual judgments made. Judgment could be described as intuitive when based on information about positive cues, pointing to a direct link between automatic learning processes and judgment independent of deliberation. Our findings led us to the intriguing conclusion that when learning about multiple cues people do not necessarily know what they are doing. Our findings support an earlier report (Rolison, Evans, Walsh, & Dennis, 2011), and offer the possibility that people utilize both implicit intuitive and explicit analytic forms of learning.

The theoretical account we propose for human judgment and multiple cue learning parallels recent theoretical developments in the category learning literature that identify multiple learning systems and distinguish controlled from automatic learning processes (Ashby & Maddox, 2005). We believe our findings point to the intuiting possibility that various forms of cue learning that are typically studied in isolation, may be understood in terms of similar underlying cognitive processes. We hope our findings inspire more research on this interesting topic.

References

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56,* 149–178.

Hammond, K. R. (1996). *Human Judgment and Social Policy*. New York: Oxford University Press

Rolison, J. J., Evans, J. St. B. T., Dennis, I., & Walsh, C. R. (2012). Dual-processes in learning and judgment: Evidence from the multiple cue probability learning paradigm. *Organizational Behavior and Human Decision Processes, 118*, 189-202.

Rolison, J. J., Evans, J. St. B. T., Walsh, C. R., & Dennis, I. (2011). The role of working memory capacity in multiple-cue probability learning. *Quarterly Journal of Experimental Psychology, 64,* 1494-1514.

Lose your Forest and You Lose Everything

Lars Sjödahl
University of Lund,
Sweden

Contact: lars.e.sjodahl@gmail.com

"Lose your forest and you lose everything" is a saying with a double imperative:

1. Look out for rebound effects when you make use of your environment (environmental resources).
2. Restore what you have exploited.

During recent decades, Sweden, like many other countries, has been struck by several devastating storms and tornados. In 2005, the storm, named Gudrun, felled over one night 75 million cubic meters of forest in the southern part of Sweden, equivalent to the total annual quantity felled under normal harvest conditions for industrial and export purposes. Forest experts are today convinced that damage of this magnitude can be prevented or reduced by initial decision when cultivating new forest stands (Blennow & Sallnäs, 2004; Olofsson & Blennow, 2012). Data from an earlier study about implementing a forest cultivation program (Sjödahl, 1984) will be presented in brief, as it has a renewed actuality, due to today's storms, causing disastrous damage to Swedish forests. Hopefully it will teach us something about a forest's initial dependence on different cultivation conditions. This study also exemplifies how Brunswik's idea about a representative situation sampling can be applied stepwise from a general domain description to a concrete item-level.

In a unifying theory for JDM-research, Hammond (1996) has suggested a Cognitive Continuum Theory (CCT) emphasizing the importance of task descriptions (see Brunswik, 1952; Juslin, 2001; Weaver & Stewart, 2012). Although tasks may vary in many respects (see Hammond, 1996, 2007, 2010) this study is confined to the following three essential task properties as they are judged to influence the actual decision process and performance quality of the cultivation tasks:

1. Tasks vary with regard to complexity. Sometimes our sampled task situations vary with regard to spatial complexity, because the ground's micro-ecology may differ from place to place, requiring an appropriate choice of plant species. Sometimes temporal complexity is the dominating, crucial property, i.e. plants have to be transported quickly, without delay to avoid damage from drought. Further, social complexity is very pronounced, as multiple actors, like

plant cultivators, transport experts, ground preparers, and land owners are involved. There is also a compromise-complexity present. Esthetic viewpoints have to be balanced against financial considerations and short-time, proximal values weighted against distal consequences.

2. Tasks vary with regard to ambiguity. To reduce the ambiguity of our task descriptions we have restricted the situation sampling by a domain specification in terms of a goal description as follows: “Our behaviour domain is defined as cultivating new forest stands, within a reasonable time period after deforestation of established forest stands, with proper regard to the ground’s productivity capacity”.
3. Tasks vary with regard to reliability of cues and number of cues. Our situation sample (n = 76) will function as cues, (= items) in three questionnaires. The items are inter-correlated, factor analyzed and the result qualitatively interpreted.

Specifying a situation sample. Based on the goal-description above, 22 main situation categories or problem-areas were formulated like 1) choice of proper planting method, 2) choice of suitable tree species, 3) understanding and implementing maps and forestry schemes, etc. These main situation categories were selected from forestry textbooks and formulated by professional forestry experts at the National Swedish Board of Forestry.

The relative need for forest workers’ and forest owners’ education on the 22 main work-categories was estimated with a 5-graded Likert scale by 29 forestry officers. The inter-judge reliability, calculated according to Winer (1962) amounts to .97, adjusted for judges’ frame of reference .98. A structured interview was then carried out with the Chief Forest Officers (n = 7) within Sweden’s growth-areas I, II and V, asking them to exemplify important work-moments within each one of the main situation categories. These interviews were recorded and content analyzed, resulting in a total of 509 task descriptions. From this pool of situations 76 different items (work situations) were selected to build a Likert scale representing each of the original 22 main situation categories. This selection carried out by forestry expertise, was based on the following guidelines:

1. The quality level of the task performance can be expected to vary between forest workers as well as between .forest owners.
2. The work situations should involve the workers judgment and decision making, i.e. not being based only on rules of thumb.
3. Each of the 22 main situation-categories should be represented in the questionnaire, preferable to equal extent.

As the aim with this study was to equip educational planners and forestry teachers with feedback information from experienced subject experts and colleagues in order to facilitate and improve their teaching planning, three questionnaires were constructed as 7-graded Likert scales, all presenting the same 76 work situations in randomized order, but addressing three different aspects.

1. The goal aspect. How important, critical, are the following 76 work operations with regard to the goal-description, given above, for cultivation of new forest stands?
2. The education, training aspect. To what extent is today’s 2-year forestry education successful with regard to the following 76 work operations?
3. The decision aspect. How often does this work situation require that the forest worker makes decisions or choices based on judgements, i.e. the tasks cannot always be managed only by rule of thumb?

The three questionnaires, representing one and the same situation sample, were sent, with proper time intervals, to all forestry teachers (n = 101) at the Swedish forestry schools. Complete responses were received from 94 teachers. The inter-judge reliability for the three questionnaires estimated according to Winer (1962) was quite high, see Table 1.

Table 1. *The inter-judge reliability of the three aspects.*

Aspects	Inter-judge reliability (n = 94)
Goal	.97 (.98)
Education	.95 (.97)
Decision	.97 (.98)

Note. In brackets: adjustment for mean differences between judges´ frames of reference

These estimates are so high that you may question the inter-independency of the single respondents. The number of forestry schools in Sweden was 20 and the number of responding forestry teachers was 94. So, we cannot exclude some "respondent-cooperation" on school-level. However inter-reliability can rise asymptotically very fast with the number of judges and it seems sometimes to be very "aspect- sensitive" (Sjödahl, 1973, 1974). To discuss the psychological dimensionality of our situation sample a descriptive, factor analysis was carried out separately for each aspect. The squared multiple correlations were chosen as communality estimates and Hotelling's iterative, principal-axes method used for extracting the factors (Pawlik, 1971). As the number of respondents (n = 94) is small relative to the number of rated work situations (n = 76) the factor analysis has been applied to the 22 main situation categories.

Consequently we calculated the average rating value for items within each of these main categories, thereby reducing the number of factor-analyzed variables to 22 average estimates. Within each aspect two factors were isolated, each presented below with corresponding three highest loaded situation categories.

Goal aspect

Factor 1

Decide point of time, when to start cultivating new forest stands	(.87)
Clearing of ground for reproduction	(.85)
Choice of appropriate plant material and tree species	(.81)

Factor 2

Handling of the single plant just before the actual planting	(.84)
The immediate planting operation	(.83)
To trim and handle the tools, necessary for planting operations	(.67)

Education aspect

Factor 1

Choice of appropriate plant material and tree species	(.82)
Transport and storage of plant material	(.80)
Choice of appropriate planting method	(.80)

Factor 2

The immediate planting operation	(.70)
Handling of the single plant just before the actual planting	(.67)
To trim and handle the tools, necessary for planting operations	(.39)

Decision aspect

Factor 1

Decide point of time (i.e. when) to start cultivating new forest stands	(.89)
To decide the need for supplementary support planting	(.82)
To decide the need for clearing the ground for reproduction measures	(.82)

Factor 2

Handling of the single plant just before the actual planting operation	(.93)
The immediate planting operation	(.88)
Transport and storage of the plants	(.85)

To get a closer look at the relationship between our three aspect-analyses, the constituents on item level included in each respective factor have been selected to form three item-sets; a goal set (G), an education set (E) and a decision set (D). The number of constituents (n) within each aspect set and within intersections and outside respective union of sets, (called specific items) are given below:
The goal set (G), n = 10, intersection with E = 3, with D = 5, specific items = 2
The education set (E), n = 12, intersection with G = 3, with D = 2, specific items 7
The decision set (D), n = 10, intersection with G = 5, with E = 2, specific 3
The intersection between the three sets G, E and D = 0

It is obvious from above, that describing a work- or task-situation from different aspects may complicate the teachers' planning, giving rise to compromises and changes of teaching conditions. Comparisons between single items' rank numbers from different aspects also show that rank differences can be extremely large. So for example is the work item/operation "to select those trees which should be left as seed-trees" ranked very high, i.e. important in the goal aspect but low in the education aspect (meaning that present teaching is insufficient), a rank difference of 30. Comparisons between the decision making and the education aspect also reveal great rank differences. The work item "to select patches of ground with local climate suitable for sowing plant seeds" is high in the goal aspect but relatively low in the education aspect, i.e. not sufficiently attended to, the rank difference being 36. Our work description so far, has mainly been in terms of work situations, i.e. subject-matter content. To make full use of these descriptions we also need an analysis in terms of psychological functions. Such a functional frame work is presented by Oberauer, Süss, Wilhelm and Wittmann (2003). The authors distinguish between three general "working memory functions", 1) simultaneous storage and processing 2) supervision 3) coordination of elements into structures, each one illustrated in turn by a task situation from our forest study as follows: 1) decide between manual, machine or chemical cleaning of planting areas, 2) detect where there is a need for complementing planting, 3) integrate knowledge about biological factors in order to decide about ground preparation.

Analyses of task situations as exemplified above may be useful for planning curricula, constructing teaching aids and training programs, and deciding about performance and achievement criteria. A qualitative interpretation of our results, i.e.,

a discussion between forestry experts with teaching experience is however, needed in order to make full use of our study-results. It is likely that such a discussion will be focused on priority problems like the following ones: What harm would there be, in the light of our goal description, to disregard the priority-ratings a) in a short-term, proximal perspective and b) in a long-term, distal perspective? What is the balance for proximal benefits like low costs, time-saving routines, minimal quality control etc. weighted against distal, negative ones like supplementary planting, additional ground clearing or preparation, changes of tree species, vulnerable to storm damage. This balance between short-term, proximal benefits and long-term, distal effects will involve compromises between values and consequences corresponding to our decision suggestions.

References

Blennow, K., & Sallnäs, O. (2004). WINDA – A system of models for assessing the probability of wind damage to forest stands within a landscape. *Ecological Modelling, 175*(1), 87-99.

Brunswik, E. (1952). The conceptual framework of psychology. In: O. Neurath, R. Carnap, & C. Morris (Eds.), *International Encyclopedia of Unified Science. Foundations of the Unity of Science, 1*(10),1-102.

Hammond, K. R. (1996). *Human judgment and social policy. Irreducible uncertainty, inevitable error, unavoidable injustice.* New York, Oxford: Oxford University Press.

Hammond, K. R. (2007). *Beyond rationality. The search for wisdom in a troubled time.* New York: Oxford University Press.

Hammond, K. R. (2010). Intuition, No!...Quasirationality, Yes!. *Psychological Inquiry, 21,* 327-337.

Juslin, P. (2001). Representative design: Cognitive science from a Brunswikian perspective. In K. R. Hammond, & T. R. Stewart (eds.). *The essential Brunswik. Beginnings, explications, applications.* (404-408). New York: Oxford University Press.

Oberauer, K., Süss, H.-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence, 31,*167-193.

Olofsson, E., & Blennow, K. (2005). Decision support for identifying spruce forest stand edges with high probability of wind damage. *Forest Ecology and Management, 207*, 1-2, 87-98.

Pawlik, K. (1971). *Dimensionen des Verhaltens.* Bern, Stuttgart, Wien: Verlag Hans Huber.

Sjödahl, L. (1973). Val av beskrivningsaspekter vid befattningsanalyser for utbildningsplanering. *Sveriges Skogsvårdsförbunds Tidskrift, 71*(3), 265-283. (In Swedish)

Sjödahl, L. (1974). Number of judges when scaling attitude items. *Scandinavian Journal of Educational Research, 18,* 183-197.

Sjödahl, L. (1984). Skogsproduktion – ett framtidsproblem. (Forestry production – a problem for the future). *Särtryck och småtryck, Nr. 482.* Lärarhögskolan, Malmö. (In Swedish).

Weaver, E. A., & Stewart, T. R. (2012). Dimensions of judgment: Factor analysis of individual differences. *Journal of Behavioral Decision Making, 25*(4), 401-413. DOI:10.1002/bdm.748.

Winer, B. J. (1962). *Statistical principles in experimental design.* New York: McGraw-Hill Book Company.

Judgment Research at Albany

Tom Stewart
University at Albany,
USA

Contact: t.stewart@albany.edu

Two papers described in last year's newsletter were published this year:

Weaver, E. A., & Stewart, T. R. (2012). Dimensions of judgment: Factor analysis of individual differences. *Behavioral Decision Making (Special Issue: Individual Differences in Decision-Making Competence), 25*(4), 402-413.

Stewart, T. R., Mumpower, J. L., & Holzworth, R. J. (2012). Learning to make selection and detection decisions: The roles of base rate and feedback. *Journal of Behavioral Decision Making, 25*(5), 522-533.

Jim, Jeryl, and I are writing up our results regarding the effect of uncertainty on learning to make selection and detection decisions.

I have been working this year with Stephen Weinberg, a behavioral economist inour faculty, on a study of individual differences in coherence competence. The only Brunswikian element in this work is that it is idiographic rather nomothetic as is most work in behavioral economics. We collected data using a within subjects design, where each subject confronts both forms of common coherence tasks (with a two week delay).

My last two students have completed their dissertations. DoSuk Lee has investigated the role of information display and uncertainty in cognitive continuum theory, building on Dunwoody's earlier study. His method differed from previous work in that he examined effects on the components of the cognitive continuum index as well as the overall index. His results indicate that there is more work needed to clarify the role of uncertainty in the theory.

Christine Muller is looking at the health care policy implications of the prioritization of medical treatments under a two stage system where policy makers prioritize treatments and then doctors prioritize patients to receive those treatments. The possible inefficiencies of such a system have not been systematically studied. Her study examined how possible differences in judgment policies between policy makers and physicians can arise and how those differences could produce unintended consequences. She found that treatment rankings can be affected by differences between policy makers and physicians (differences that could easily be overlooked in implementing health care policy), but that the effect may be small unless differences in the importance of the prioritization criteria are quite large.

Evidence-Based Practice in Education Requires (Quasi-)Representative Designs

Christof Wecker
Ludwig-Maximilians-Universität München,
Germany

Contact: christof.wecker@psy.lmu.de

Often the practitioner side is blamed for the lack of evidence-based practice in education (e. g. Fox, 2003, p. 95 f.). Educators are sometimes pictured as being ignorant of research findings or unwilling to take the trouble to apply them in their work. But does research actually deliver what is needed to take rational decisions supported by strong empirical evidence?

Rational decisions in education would be based, amongst other things, on general statements about regularities that link a certain type of instructional strategy, e. g. the use of graphical representations in instruction, to certain educationally relevant outcomes, such as the acquisition of knowledge or skills, a decrease or increase of interest or a more or less favourable self-concept (Wecker, in press). The charge of lacking rationality or responsibility cannot be brought against practitioners failing to apply a certain type of instructional strategy unless the generality of such a general statement has been firmly established. Drawing upon Brunswikian ideas, in a recent publication (Wecker, in press) I dealt with this issue in the context of the justification of recommendations for practice. In the following, I will summarize the main argument.

It is important to note first that a statement about a regularity concerning the probabilistic connection between an instructional strategy and an educational outcome remains a general claim, irrespective of the number of qualifications that restrict it to certain age groups, school subjects, and the like. The reason for this is that there will always be an infinite number of possible in-stances of any type of instructional strategy in question, which may differ in all kinds of – “relevant” or “irrelevant” – detail.

The widely accepted practice of testing a general statement about a “population” in the social sciences involves drawing an appropriate kind of sample of persons and applying appropriate inferential statistical tests to the data collected from it. Depending mainly on the appropriateness of the sampling technique, this approach is more or less appropriate for providing evidence for the generality of a statement with respect to a population of persons. At the same time, as it stands, this approach is entirely inappropriate for providing any evidence for the generality of a statement in any other respect, such as across school subjects, instructors, or institutions, and in particular, the possible instances of a type of instructional strategy. Typically, intervention studies in educational research are conducted within one content area (for good reasons), and in many studies a standardized version of the treatment is administered with the help of a computer (often for not so good reasons). This amounts to a sample size of 1 with respect to domains and possible instances of a type of instructional strategy. Hence, studies of this kind do not provide any evidence

for the generality of a statement about a regularity concerning the connection between a certain type of instructional strategy and an educationally relevant outcome.

This is a severe problem for two reasons: First, as argued above, it is a prerequisite for evidence-based practice that the generality of statements about regularities is firmly established with respect to aspects beyond the population of persons it is supposed to apply to. Second, the theoretical claims made in scientific articles are hardly ever explicitly or implicitly restricted to the particular instances of a type of instructional strategy implemented in the study, despite the fact that authors frequently switch to the past tense in the discussion section as if they were not interested in any broader conclusions beyond their sample of participants and research setting.

The solution for this problem is quite straightforward and entirely analogous to the issue of generality with respect to a population of persons: Evidence for generality with respect to aspects such as possible instances of a type of instructional strategy, schools subjects, instructors, and institutions, requires drawing an appropriate kind of sample from these “universes” (Brunswik, 1955, pp. 198; 202; 1956, p. 37; Cronbach, Gleser, Nanda & Najaratnam, 1972, p. 18; Snow, 1974, p. 272) and applying appropriate inferential statistical tests to the data collected from it. (The term “universe” is clearly more appropriate than the term “population” because it does not invite the fallacy that a sample can only consist of persons rather than of “tuples” of persons, instances of a certain type of instructional strategy, and further aspects of the situation.) This amounts to nothing more than a call for representative (Brunswik, 1955, pp. 198; 202; 1956, p. 37) or quasi-representative (Snow, 1974, pp. 271; 273) designs in educational research.

Certainly not every single study can implement a design that is representative with respect to every aspect of generality of the statement it is designed to scrutinize. After all, the statement can fail this test even in a more restricted systematic design with aspects held constant. However, as an area of research becomes mature and researchers proceed to voice recommendations for practice, such studies become imperative.

Two particular obstacles for this kind of studies in educational research are the difficulties (1) to sample instances of a certain type of instructional strategy and (2) to compare important educational outcomes such as knowledge and skills across content domains, which seem to prevent studies that sample across content domains. I have sketched a solution for these obstacles (Wecker, in press), which involves drawing a sample of instructors or instructional designers who are assigned different topics and each of which is asked to develop two versions (corresponding to the treatment and control conditions) of an instructional unit. Then they teach their units to samples of learners. The data from this design can be analysed by means of multi-level modelling, taking into account the nesting of participants within instructors and topics. This approach allows for estimating the variability of the effectiveness associated with different instances of the type of instructional strategy implemented by different instructors, and thereby provides important information concerning generality.

Issues such as the ones discussed here are often conceived of exclusively as a matter of so-called “external validity”, which is typically regarded as less crucial than the elimination of alternative potential causes of effects (Campbell & Stanley, 1966, p. 5). This view misses an important point: Experimental manipulations can be described in different ways, using terms with different extensions. For example, a

single treatment might fulfill the defining criteria of “feedback” and of “criticism” at the same time. If only one version of the treatment is implemented in a study, both terms may be equally justified post hoc. Only if a set of instances of the treatment has been sampled based upon a general specification of the type of treatment in advance, a decision about the actual cause of an observed effect can be justified. Hence, the issues discussed also touch upon the issue of finding the real causes of an effect – which Brunswik seems to have been aware of in contrast to current mainstream methodology. There is much to learn from his ideas that can help us in providing a robust research base for evidence-based practice in education.

References

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*(3), 193-217.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles.* New York: Wiley.

Fox, M. (2003). Opening Pandora's Box: Evidence-based practice for educational psychologists. *Educational Psychology in Practice, 19*(2), 91-102.

Snow, R. E. (1974). Representative and quasi-representative designs for research on teaching. *Review of Educational Research, 44,* 265-293.

Wecker, C. (in press). How to support prescriptive statements by empirical research: Some missing parts. *Educational Psychology Review*. doi:10.1007/s10648-012-9208-9

A Brunswikian Lens Model of Personality Impressions from Facebook Profiles

Aaron C. Weidman
University of British Columbia, Vancouver, BC, Canada

Nicholas S. Holtzman
Washington University in St. Louis, MO, USA

Contact: acweidman@psych.ubc.ca, nick.holtzman@gmail.com

The sharp increase in online social networking (OSN) use over the past decade has led to questions about how personality is revealed in digital mediums, such as OSN profiles. Researchers have begun to use Brunswik's (1956) lens models to examine the trait-relevant cues that help observers gain insight into the personality of OSN profile owners, such as the number of photos posted by users and the amount of personal information revealed in profiles (Marcus, Machilek, & Schütz, 2006). When using such cues, observers tend to agree with each other and achieve a substantial level of accuracy of their impressions (Back, Stopfer, Vazire et al., 2010; Vazire & Gosling, 2004). Building on this work, Gosling, Augustine, Vazire, Holtzman, and Gaddis (2011) ran two studies to examine personality judgments from Facebook profiles. They employed a lens model to identify whether judges agree when forming impressions (i.e., consensus), the extent to which judges are good judges (i.e., accuracy), which cues are meaningful indicators of personality (i.e., cue validity), and which cues judges use to form impressions (i.e., cue utility).

To identify some of the valid personality cues in Facebook profiles, in Study 1 young adults completed a questionnaire measuring the Big 5 personality traits and reported their engagement in a range of Facebook behaviors (e.g., number of Facebook friends, number of photos). The relationship between each behavior and participants' trait levels was then examined. The results suggested that extraverted individuals leave traces of their sociable, outgoing, and gregarious behavior on Facebook; extraverted participants reported having more Facebook friends, viewing and commenting on other users' walls more frequently, adding more photographs, and spending more time on Facebook. Other personality traits showed only small and infrequent prediction of Facebook behaviors (e.g., openness predicted frequently changing one's profile picture). Thus, Study 1 indicated that most of the behavioral information on OSNs concerns the owners' extraversion; Facebook profiles are likely a gold mine for extracting extraversion cues.

Study 1 relied on self-reported Facebook behavior, however, and as a result may not have represented users' actual online behavior. The authors conducted a second study in which young adults consented to having their Facebook profiles coded by unacquainted observers for behaviors (e.g., number of networks joined, indicating sociality). Participants' self-reported personality was combined with personality ratings provided by well-acquainted peers to form a criterion personality score.

A traditional Brunswikian lens model analysis was employed to examine judgments of profile owners' personalities. First, the authors found that observers formed consensus judgments of profile owners on each trait, meaning that unacquainted raters largely agreed about profile owners' personalities. This finding indicates that Facebook profiles give off a consistent impression about their owners across a range of observers. Second, accuracy was measured by correlating the average personality rating provided by the unacquainted observers with the criterion personality score. Observers achieved a high degree of accuracy in judging extraversion and openness, modest accuracy in judging conscientiousness and agreeableness, and no accuracy when judging neuroticism. In contrast to Study 1, these findings suggest that Facebook profiles give off diagnostic information about one's personality, a question that cannot be examined merely by asking individuals to self-report their personalities. It appears that observers can know a fair amount about someone's personality — even without meeting them — just by briefly viewing their Facebook profile.

Next, the authors examined the valid cues provided in Facebook profiles by correlating the criterion of owners' personality with the coded score on each Facebook behavior. Consistent with Study 1, results showed that cues of extraversion appear all over Facebook profiles; highly extraverted profile owners had joined many groups and networks and had large numbers of friends, photos, and wall posts. Profiles provided a few valid cues of openness, such as a large number of friends and networks, though these cues did not sufficiently distinguish highly open from highly extraverted Facebook users. These results support theoretical accounts of extraversion and openness; extraverts are highly sociable, outgoing, and gregarious, and thus we would expect them to make many friends and have vibrant social lives on Facebook. Similarly, open individuals crave a variety of social experiences, and not surprisingly seek to make many acquaintances and engage with a range of social groups, organizations, and activities. Importantly, these findings build upon those of Study 1 by suggesting that extraverts don't just claim to act sociable online, they in fact do act sociable online.

Even if Facebook profiles do provide valid personality cues, observers must use those valid cues to achieve accurate judgments. The authors examined cue utility by correlating the average personality rating provided by the unacquainted observers with the coded score on each Facebook behavior. Results showed that observers picked up on all the valid cues of extraversion; observers saw high levels of extraversion in owners who had many friends, photos, and wall posts, as well as those who belonged to many groups and networks. Observers also inferred profile owners' openness from a large number of photos and groups.

These findings suggest that extraversion is the most visible trait in a Facebook profile, as unacquainted observers correctly perceived behavioral cues of extraversion, such as having many friends, and used these cues to form impressions. Despite the intuitive notion that the internet is a medium through which individuals with fewer social skills feel that they can comfortably compensate for relationship difficulties offline (Weidman et al., 2012), the extraversion finding serves as a reminder that the primary beneficiaries of OSNs are outgoing, gregarious individuals with many existing friendships. Consistent with this theory, recent research has shown that low self-esteem individuals' online self-disclosures (i.e., Facebook status updates) tend to contain unpleasant emotional information, and that this may repel potential friends and acquaintances (Forest & Wood, 2012). This

finding also reflects the general rule of thumb that online personality reflects offline personality (Back et al., 2010).

Why weren't agreeableness, conscientiousness, and neuroticism associated with any utilized or valid Facebook cues? These traits may be relatively internal and unobservable on Facebook, and thus may not be manifested in specific cues (John & Robins, 1993; Vazire, 2010). For example, a neurotic profile owner who posts on a friend's wall might worry about whether or not that friend will actually read and respond to the wall post, but these ruminations will not appear in any visible way to the outside observer. Similarly, traits such as agreeableness and conscientiousness might manifest in very subtle linguistic cues not measured in this study, and such subtle cues might be allowing observers to make the moderately accurate judgments found in this study; a few speculative possibilities are that agreeable profile owners may respond to wall posts with more positively worded messages, whereas conscientious profile owners may respond with better punctuation.

In conclusion, Gosling and colleagues' (2011) studies add to a growing body of research examining personality impressions from online social networking sites and other virtual environments (see Wilson, Gosling, & Graham, 2012, for a review) that represents a novel application of the Brunswikian lens model to understanding the processes behind everyday personality judgments. The lens model has helped psychologists understand which observable, trait-relevant cues are the mechanisms by which impressions are made across many contexts, including Facebook profiles (Vazire & Gosling, 2004; Back et al., 2010), blogs (Yarkoni, 2010), email addresses (Back, Schmuckle, & Egloff, 2008), and even text messages (Holtgraves, 2011). It appears that, for example, if your friend is a gregarious goof in everyday life, then in the digital world your friend probably seems gregarious and goofy.

References

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmuckle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science, 21,* 372-374.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, CA: University of California Press.

Forest, A. L., & Wood, J. V. (2012). When social networking is not working: Individuals with low self-esteem recognize but do not reap the benefits of self-disclosure on Facebook. *Psychological Science, 23,* 295-302.

Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N. S., & Gaddis, S. (2011). Manifestations of personality in online social networks: Self-reported Facebook-related behaviors and observable profile information. *Cyberpsychology, Behavior, and Social Networking, 14,* 483-488.

Holtgraves, T. (2011). Text messaging, personality, and the social context. *Journal of Research in Personality, 45,* 92-99.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61*, 521-551.

Marcus, B., Machilek, F., & Schütz, A. (2006). Personality in cyberspace: Personal websites as media for personality expressions and impressions. *Journal of Personality and Social Psychology, 90,* 1014-1031.

Valkenburg, P. M., & Peter, J. (2009). Social consequences of the internet for adolescents: A decade of research. *Current Directions in Psychological Science, 18,* 1-5.

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology, 98,* 281-300.

Vazire, S., & Gosling, S. D. (2004). E-perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology, 87,* 123-132.

Weidman, A. C., Fernandez, K. C., Levinson, C. A., Augustine, A. A., Larsen, R. J., & Rodebaugh, T. L. (2012). Compensatory internet use among individuals higher in social anxiety and its implications for well-being. *Personality and Individual Differences, 53,* 191-195.

Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A review of Facebook research in the social sciences. *Perspectives on Psychological Science, 7,* 203-220.
Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality, 44*, 663-673.

Brunswik-Symmetry

Werner W. Wittmann
Canada Otto-Selz-Institute for Applied Psychology,
University of Mannheim,
Germany

Contact: wittmann@tnt.psychologie.uni-mannheim.de

I was invited by Professor Klaus Scherer at the University of Geneva to give a talk entitled: “Brunswik-symmetry, a golden key concept for a successful psychological science”. I took the opportunity to summarize the research my coworkers and I have done so far using this concept. The presentation contained research on solutions to the Mischel-Epstein debate about the role of personality and situation in predicting behavior. The relation between working memory constructs and psychometric intelligence was taken up. Further, meta-analyses on psychotherapy outcome research were presented (illustrated). Performance in complex computer based problem solving scenarios was illustrated and discussed in relation to intelligence, knowledge and working memory. Theory-derived suppressor-principles were highlighted as means to increase predictability by preventing the subject from looking in the wrong direction. The dummy variable used to contrast treatment and control in experimental research is often very unreliable and attenuates effect sizes. It is shown how these concepts are to be included in the application of the lens model equation and Brunswik’s symmetry concepts.

The presentation can be downloaded at:
http://www.psychologie.uni-mannheim.de/brunswik-symmetry/

Klaus is one of the most prominent researchers in the area of emotion and a long time fan of Brunswikian ideas to be seen in many of his publications. His extensive publication list can be found at:
http://www.affective-sciences.org/publications/author/scherer
There you will find papers where he applies Brunswikian concepts.
At the Mannheim location are additional slides of a presentation given at the American Evaluation Association conference in Minneapolis, MN in October 2012. The theme of the conference was “Evaluation in Complex Ecologies, Relationships, Responsibilities and Relevance” concepts, which are all concerned with Brunswik’s representative design. Moreover different pros and cons dealing with the classical randomized control group design in complex ecologies were discussed.

Notices

Call for Papers

Julian N. Marewski & Ulrich Hoffrage
Faculty of Business and Economics (HEC),
University of Lausanne,
Switzerland

Contact: julian.marewski@unil.ch

The Journal of Applied Research in Memory and Cognition (JARMAC) will publish a special issue on "Modeling and Aiding Intuitions in Organizational Decision Making", edited by Julian N. Marewski and Ulrich Hoffrage.

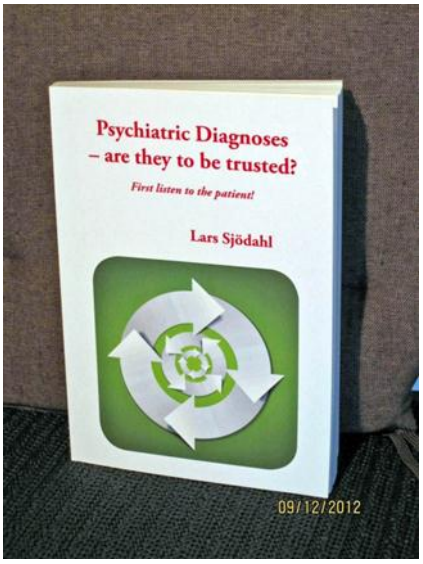
Interested contributors are requested to contact Julian Marewski and Ulrich Hoffrage (julian.marewski@unil.ch, ulrich.hoffrage@unil.ch; for more information about the guest editors, see www.modeling-adaptive-cognition.org) and to submit, as a preliminary step, a summary of the intended contribution (about 200 words). Each summary till be evaluated by the guest editors in terms of the intended contribution's scope and suitability for the special issue. Summaries that are submitted prior to December 31st will be given full consideration for the special issue: summaries that are submitted on a later date will also be considered: however, full consideration of late summaries will only be guaranteed as long as projected number of intended contributions does not exceed the available journal space. The deadline for submitting full papers is October 15th, 2013. Submitted papers will be reviewed within 4 weeks after their reception.

New Book

Psychiatric Diagnoses – Are they to be Trusted? First Listen to the Patient!

Lars Sjödahl
University of Lund,
Sweden

Contact: lars.e.sjodahl@gmail.com



Psychiatric diagnoses - are they to be trusted? presents a new critique of today's psychiatric diagnostics and treatment. The book is based on 20 recovered psychiatric patients' own evaluation of the content of their patient charts. The manifold single case illustrations are discussed against the background of today's psychological and psychiatric research.

Patients are no longer satisfied with only being talked about by somebody else. They want to speak with their own voice and want a more holistic perspective of their situation. This recent change has opened up for improved transparency within psychiatric health care. Today's psychiatric patients want answers to the following questions:

1. What does my diagnosis mean? Would I get a different diagnosis from another doctor?
2. Are the notes in my patient chart true in relation to the reality I know and could be checked independently by me?
3. Which alternative treatments can be offered?
4. How is the treatment prescribed for me related to my diagnosis?
5. How do I distinguish between my symptoms and the side-effects from prescribed medication?

What the single patient says and what the doctor has put in the patient's chart are not always compatible. Discrepancies of this kind seem often to be indications of serious flaws and inconsistencies, which on closer inspection reveal more profound weaknesses of personal or organizational nature. It is always the patient who has to bear the serious consequences of these discrepancies.

The author's argument is borne out by copious references to leading researchers within the fields of Psychology, Psychiatry and Sociology.

Lars Sjödahl worked at the University of Lund, Sweden, as lecturer and researcher in occupational and educational psychology. He has also considerable experience from working with personnel administration within industry and healthcare. He is the author of a number of books in Swedish dealing with his research and since 2007 has been the editor of the annual Brunswik Society Newsletter http://www.brunswik.org/newsletters/index.html

lars.e.sjodahl@gmail.com
ISBN 978-91-7465-4134
This book can be ordered from
www.books-on-demand.com

To get information about future Newsletters, you are welcome to join (free of charge) the email list of the Brunswik Society at:

http://www.brunswik.org/