The Brunswik Society

Newsletter

Vol. 35, Nov. 2020

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

http://www.brunswik.org/

ISSN 2296-9926

Table of Contents

Foreword

Kaufmann, E., & Hamm, R. M..... 3

Contributions

Broomell, S. B.
Using a Bifocal Lens Model to Identify Gobal-Local Incompatibility..... 5

Chacon, A., Herrera, N., Kausel, E., & Reyes, T., Kaufmann, E.
Computerized Algorithm Advice from a Brunswikian Perspective..... 8

Holleman, G. A.
Ecological Validity and the “Real World" in the Field of Social Attention..... 9

Kihlstrom, J. F.
Three Meanings of “Ecological Validity"..... 11

Ladinig, T. B., Dhir, K. S., & Vastag, G.
Sensemaking Support System (S^3) for Manufacturing Process Improvement..... 13

Prätor, S., & Guéridon, M.
Clinical Judgments of Recidivism by Prison Staff and Official Recidivism Rates of Female Prisoners: An Application of Brunswik’s Lens Model in a Women’s Prison in Lower Saxony, Germany..... 16

Unkelbach, C., Mierop, A., & Corneille, O.
A Brunswikian View on Judging Fake News in Social Media Communications..... 19

Yu, M. C., & Kuncel, N. R.
Pushing the Limits for Judgmental Consistency: Comparing Random Weighting Schemes with Expert Judgments..... 22

Luan, S., Schooler, L. J., & Tan, J. H.
Improving Judgment Accuracy by Sequential Adjustment..... 25

Agenda

26th International (Virtual) Meeting of the Brunswik Society..... 29

Imprint
The Brunswik Society Newsletter, ISSN 2296-9926
is published yearly
Editorial staff: Kaufmann, E., & Hamm, R. M.

Contact:
The Brunswik Society Newsletter
c/o University of Zurich
Institute of Education, University of Zurich,
Kantonsschulstr. 3
CH-8006 Zurich
Switzerland

Foreword

Many thanks to all authors for their contributions!

We are very pleased to present the 2020 Brunswik Society Newsletter.

Due to the COVID-19 challenges we all faced this year worldwide, we are extremely proud that the Brunswik Society successfully adapted to the circumstances by initiating the first virtual Brunswik Society meeting. We also consider this first virtual Brunswik Society meeting within this Newsletter. The complete program and contribution abstracts are presented. The two-day meeting covers contributions on judgment and decision making under the label “Expert judgment and the lens model” as well as methodological contributions subsumed under the label “Research design and adaptive cognition”. The two main emphases show also the richness of Brunswik’s and Hammond’s legacy.

The current version of the newsletter starts with Broomell’s contribution introducing a new version of a bifocal lens model and applying it to judgments of Tornado season.

This year, Brunswik's concept of ecological validity (Hammond’s 1988) came up and was revived within a discussion of Perspectives and Psychological Science – to catch up with this discussion, we included two contributions on that topic. Kihlstrom’s contribution considers also personal historical references which is greatly recommended to read. Holleman’s contribution and her presentation at the next Brunswik Society meeting will certainly also relaunch the discussion on ecological validity within the Brunswik Society.

Unkelbach and colleagues' contribution brings Brunswik research to the digital world applying the Brunswikian lens to the social media ecology. Another application of the lens model framework is given in the contribution by Prätor and Guéridon in predicting the recidivism of female prisoners by prison staff members. Ladinig and colleagues applied the Brunswikian approach in the development of a sensemaking support system to help operations managers find mutually acceptable solutions based on their judgments.

Several contributions within this newsletter focus especially on improving judgment accuracy. Yu and Kuncel compared random weighted schemes with expert judgments showing the importance of judgmental consistency to increase judgment accuracy. Luan and colleagues demonstrated that presenting cues sequentially instead of simultaneously increases judgment accuracy, too. Whether computerized algorithm advice has the power to increase judgment accuracy is the topic in the overview by Chacon and colleagues.

We hope the richness of this newsletter and the up-coming next Brunswik Society meeting will inspiring you to conduct Brunswik-Hammond research or/and become or stay an active member of the Brunswik Society.

Sincerely,
Esther Kaufmann and Robert M. Hamm

Thank you to Tom Stewart, the webmaster of the Brunswik Society, for providing web access to the Newsletter.

If you're interested in supporting the editorial team of the Brunswik Society Newsletter and to be involved in the next Brunswik Society Newsletter let us know by email (esther.kaufmann@gmx.ch). Thank you in advance for your support.

Contributions

Using a Bifocal Lens Model to Identify Global-Local Incompatibility

Stephen B. Broomell
Carnegie Mellon University, USA

Contact: broomell@cmu.edu

I propose a model of judgments about large scale, global variables that exist beyond the perception of a decision maker (DM). Such variables include the global average temperature of earth, or the likelihood of tornadoes for a particular region. Generally, I define a global variable to be any variable that a DM attempts to judge that cannot be directly observed, similar to the notion of a population level parameter in statistics.

The key element of this judgment context is that a DM must integrate multiple cues (with uncertainty about their validity and reliability) to form a judgment. Brunswik's (1952) lens model depicts how the environment facilitates accurate judgment of distal variables. Expanding the lens model to global variables, I define a global lens as the full collection of cues that can, in theory, be used to measure the state of a global variable. I define the local lens as a subset of cues and observations that DMs actually use to form a judgment (constrained by the limits of their own perception and memory). Adding the local lens to the global lens turns the traditional lens model into a bifocal lens (as depicted in Figure 2 of Broomell, 2020). This model differs from Beckstead (2017), who also presents a bifocal lens model.

For example, consider judging the average temperature of a U.S. state at one particular moment in time. The global set of cues would be measurements of temperature at every location in the state. An individual resident of the state is constrained to observing temperatures in their geographic location. Incompatibility results from the high agreement between localized temperatures failing to make clear the variability of observations spanning an entire state. Judgments of the average temperature for the state will be highly unreliable, but confidence in such judgments will be increased by the consistency of the local environment (see Section 4 of Broomell, 2020 for an empirical demonstration).

A typical regression-based lens model analysis focuses on achievement by determining whether the weight of each cue in forming a judgment matches the true relationship between each cue and the global variable (Tucker, 1964). The approach adopted in Broomell (2020) is based on classical test theory, focusing instead on the reliability of the local perspective for measuring the true state of a global variable. In other words, I recast Tucker's (1964) formulation of the lens model as a measurement problem. If cues appear reliable locally, a DM may also assume they are reliable globally. I theorize

that a DM's confidence in his or her judgment of a global variable is based on the internal consistency of the local cues (similar to estimating Cronbach's α).

My focus on reliability pinpoints a problem with global variables that goes beyond ecological validity (i.e., finding the most valid cues), as a problem of ecological reliability. Even if a DM relies on the most ecologically valid observations, the noise in local information can be so great that it cannot reveal the current state of the global variable. In such a context, DMs will need more cues, from different vantage points, to average out the measurement error in order to reach a reliable judgment. Incompatibility arises when the local environment appears internally consistent despite its lack of reliability for measuring a global variable.

This approach directly models the information content of observations, revealing informational bottle-necks in the environment that will hinder the ability of any cognitive process to generate accurate judgments, regardless of its ecological validity. As such, this framework bridges the study of judgment across diverse decision contexts by identifying the basic statistical properties that make environments appear more informative than they truly are.

Shafir (1995) applied stimulus-response compatibility to understand how the context of information can alter judgments. Certain stimuli can enhance certain compatible responses, and impede certain incompatible responses. Global-local incompatibility proposes that a collection of internally consistent cues will enhance a DM's reliance on such cues for judgment. Therefore, judgments based on these cues are mentally incompatible with treating them as noisy and uninformative, which is often the case for global variables.

An Application: Judgments of Tornado Season

The Southeastern region of the United States experiences high numbers of tornado fatalities. Knowledge about tornado likelihood may be especially important in the Southeast because this region lacks a single, traditional tornado season. Global-local incompatibility suggests that judgments of tornado likelihood based on local observations are susceptible to inflated perceptions of reliability.

Broomell et al. (2020) investigated how Southeast residents judge the likelihood of tornadoes (a) during different seasons of the year, (b) at different times of day, and (c) from different types of storm systems. Overall, there was a lot of variability in responses provided by this sample, revealing a lack of consensus about when tornadoes are more likely to occur. Relative to an expert sample, Southeast residents on average overestimated the likelihood of tornadoes in summer, underestimated their likelihood in winter, and underestimated tornado likelihood at night.

Expert judgments of tornado likelihood in the Southeast are based on large data sets documenting all the known tornadoes in the region. Residents of the Southeast have local experiences that can dramatically differ from what the tornado database shows. These differences stem from a global-local incompatibility due to the low base rate (and concentrated impact) of tornadoes.

Conclusions

Brunswik brought attention to how the environment can shape perceptions and judgments. Building on this perspective, I propose that environments shape confidence in judgments. In many situations, consistency in decision environments is informative of the reliability of our judgments. However, in many societally important contexts that involve global variables, environmental consistency is incompatible with our judgment processes. This theory can help explain a lack of consensus about global variables (such as climate change) and provide insights into approaches for helping to calibrate confidence in judgments about such issues.

For the full article, please see:

Broomell, S. B. (2020). Global–Local incompatibility: The misperception of reliability in judgment regarding global variables. *Cognitive Science, 44*(4), e12831. https://doi.org/10.1111/cogs.12831

References:

Beckstead, J. W. (2017). The bifocal lens model and equation: Examining the linkage between clinical judgments and decisions. *Medical Decision Making, 37*(1), 35–45. https://doi.org/10.1177/0272989X16674196

Broomell, S. B. (2020). Global–Local incompatibility: The misperception of reliability in judgment regarding global variables. *Cognitive Science, 44*(4), e12831. https://doi.org/10.1111/cogs.12831

Broomell, S. B., Wong-Parodi, G., Morss, R. E., & Demuth, J. L. (2020). Do we know our own tornado season? A psychological investigation of perceived tornado likelihood in the southeast united states. *Weather, Climate, and Society, 12*(4), 771–788. https://doi.org/10.1175/WCAS-D-20-0030.1

Brunswik, E. (1952). The conceptual framework of psychology. *International Encyclopedia of Unified Science*, *1*(10), University of Chicago Press.

Shafir, E. (1995). Compatibility in cognition and decision. *Psychology of Learning and Motivation*, *32*, 247–274. https://doi.org/10.1016/S0079-7421(08)60312-6

Tucker, L. R. (1964). A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd. *Psychological Review, 71*(6), 528–530. https://doi.org/10.1037/h0047061

Computerized Algorithm Advice from a Brunswikian Perspective

Alvaro Chacon, Nicolas Herrera, Edgar E. Kausel, Tomas Reyes
Pontificia Universidad Católica de Chile, Chile

Esther Kaufmann
University of Zurich, Switzerland/University of Konstanz, Germany

Contact: achacon@uc.cl

Within the Brunswikian framework, the building of expert models is often done through the use of the indexes of the lens-model equation (see e.g., Tucker, 1964). These models are provided to support judgment and decision makers with advice (see Karelaia & Hogarth, 2008; Kaufmann & Wittmann, 2016).

Due to the increase of data by digitalization, we expect that the need for advice by digital expert models is increasing. However, it is unclear if such advice is accepted by the judgment and decision-maker – especially if he/she is an expert in the field in which a judgment or decision must be made, as is the case with, for example, a physician or a teacher.

Previous research mostly shows an algorithm aversion meaning that expert models' advice is not preferred over human advice (see e.g., Burton et al., 2020; Kaufmann & Budescu, 2020). However, up to now, additional knowledge on tasks or judges, or even an overview thereof, is missing. Hence, we are working on an overview of the acceptance of expert models guided by Brunswik's lens model framework considering in detail task- and individual characteristics as well as methodological study aspects.

Our current review has identified 40 studies, including 115 sub-studies (e.g., where participants perform a task under certain circumstances) and 32,529 participants from algorithmic advice acceptance literature. Among other issues, our review suggests that future studies and evaluations on the acceptance of expert model advice consider Hammond's Cognitive Continuum Theory. In this area, we see the potential for the evaluation of algorithm advice. In addition, we suggest an overview of different theories on the acceptance of technology tools to check the fruitfulness of Hammond's Cognitive Continuum Theory in detail.

We provide additional material for interested readers and greatly welcome comments on or critique of our research agenda.

References:

Burton, J. W., Stein, M., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making, 33*(2), 220–239. https://doi.org/10.1002/bdm.2155

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin, 134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Kaufmann, E., & Budescu, D. V. (2020). Do teachers consider advice? On the acceptance of computerized expert models. *Journal of Educational Measurement, 57*(2), 311–342.

https://doi.org/10.1111/jedm.12251
Kaufmann, E., & Wittmann, W. W. (2016). The success of linear bootstrapping models: Decision domain-expertise-, and criterion-specific meta-analysis. *PLoS ONE, 11*(6), e0157914. https://doi.org/10.1371/journal.pone.0157914
Tucker, L. R. (1964). A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd. *Psychological Review, 71*(6), 528–530. https://doi.org/10.1037/h0047061

Ecological Validity and the "Real World" in the Field of Social Attention

Gijs A. Holleman
Experimental Psychology, Helmholtz Institute & Developmental Psychology, Utrecht University, the Netherlands

Contact: g.a.holleman@uu.nl

My colleagues and I are interested in gaze behavior and visual attention during various human interactions, such as face-to-face communication (Hessels et al., 2019; Holleman et al., 2020), or potential interactions in human crowds (Hessels et al., 2020). The study of the relations between gaze, attention, and the social environment is often referred to by the term social attention. In this field, there has been a recent surge (2006-present) in papers advocating for more real-world research (Kingstone, 2009; Risko et al., 2016; Smilek et al., 2006). The main argument is that in order to overcome some of the limitations of lab-based experiments, researchers need to design experiments with "more" or "higher" ecological validity in order to better approximate real-life behavior (Osborne-Crowley, 2020; Risko et al., 2012; Shamay-Tsoory & Mendelsohn, 2019). This argument is certainly not new, nor has this discussion been limited to the field of social attention for that matter, as pointed out in detail by Hammond (1998).

Being familiar with Brunswik's and Hammond's work, we have noticed that discussions within the field of social attention about the 'real world or the lab'-dilemma have been severely hampered by a lack of terminological precision, as well as by misleading and intuitive assumptions about how one could (or should) achieve 'more ecological validity' to better approximate 'real life' behavior. We wrote two articles with the aim of addressing the superficiality of these discussions, as well as educating a new generation of researchers about the history and problems associated with the popular usage of the term ecological validity. In the first article (Holleman et al., 2020a), we provide a brief historical overview of the 'real-world or the lab'-dilemma, we discuss several problems and assumptions associated with the current usage of the term 'ecological validity', and we inform researchers about Brunswik's original definition. Also, we highlight Hammond's (1998) critique of empty phrases such as 'real life" and the 'real world' and apply this reasoning to the field of social attention. In a second article (Holleman et al., 2020b), we commented on a recently published review by Shamay-Tsoory and Mendelsohn (2019) which again advocated for experimental paradigms possessing "high ecological validity" (p. 851). Although the authors were familiar with Brunswik as the originator of "ecological validity" and "representative design", they did not seem to utilize its theoretical basis to their advantage. We set out to correct common misunderstandings about these concepts,

and we aimed to show how Brunswik's method of representative design may contribute to Shamay-Tsoory & Mendelsohn's proposed ecological approach.

Although we realize that researchers in the field of social attention are perhaps not likely to change their terminology or conceptual frameworks overnight, our primary goal is to raise the level of discussion in this field by introducing researchers to some of Brunswik's principal ideas and concepts.

References:

Hammond, K. R. (1998). *Ecological validity: Then and now.* Retrieved from: http://www.brunswik.org/notes/essay2.html

Hessels, R. S., Holleman, G. A., Kingstone, A., Hooge, I. T. C., & Kemner, C. (2019). Gaze allocation in face-to-face communication is affected primarily by task structure and social context, not stimulus-driven factors. *Cognition*, *184*, 28–43. https://doi.org/10.1016/j.cognition.2018.12.005

Hessels, R. S., van Doorn, A. J., Benjamins, J. S., Holleman, G. A., & Hooge, I. T. C. (2020). Task-related gaze control in human crowd navigation. *Attention, Perception, & Psychophysics*, *82*, 2482–2501. https://doi.org/10.3758/s13414-019-01952-9

Holleman, G. A., Hessels, R. S., Kemner, C., & Hooge, I. T. C. (2020). Implying social interaction and its influence on gaze behavior to the eyes. *PLoS ONE*, *15*(2). https://doi.org/10.1371/journal.pone.0229203

Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020a). The 'real-world approach' and its problems: A critique of the term ecological validity. *Frontiers in Psychology*, *11*, 721. https://doi.org/10.3389/fpsyg.2020.00721

Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020b). The reality of "real-life" neuroscience: A commentary on Shamay-Tsoory and Mendelsohn (2019). *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691620917354

Kingstone, A. (2009). Taking a real look at social attention. *Current Opinion in Neurobiology*, *19*(1), 52–56. https://doi.org/10.1016/j.conb.2009.05.004

Osborne-Crowley, K. (2020). Social cognition in the real world: Reconnecting the study of social cognition with social reality. *Review of General Psychology*, *24*(2), 144–158. https://doi.org/10.1177/1089268020906483

Risko, E. F., Laidlaw, K., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, *6*, 143. https://doi.org/10.3389/fnhum.2012.00143

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. *Current Directions in Psychological Science, 25*(1), 70–74. https://doi.org/10.1177/0963721415617806

Shamay-Tsoory, S. G., & Mendelsohn, A. (2019). Real-Life neuroscience: An ecological approach to brain and behavior research. *Perspectives on Psychological Science*, *14*(5), 841–859. https://doi.org/10.1177/1745691619856350

Smilek, D., Birmingham, E., Cameron, D., Bischof, W., & Kingstone, A. (2006). Cognitive ethology and exploring attention in real-world scenes. *Brain Research*, *1080*(1), 101–119. https://doi.org/10.1016/j.brainres.2005.12.090

Three Meanings of "Ecological Validity"

John F. Kihlstrom
University of California, Berkeley, USA

Contact: jfkihlstrom@berkeley.edu

In a paper widely circulated in *samizdat* and now available on the Brunswik Society website, Hammond (1998) criticized social psychologists and others for misusing the concept of "ecological validity". He correctly stated that in, coining that term, Brunswik was referring to the validity of the cues available for perception and judgment, not the validity of experiments. The issue came up again recently in the pages of *Perspectives on Psychological Science* (*PPS*), when Simone Shamay-Tsoory and Avi Mendelsohn (2019) published an article advocating for the use of more "ecologically valid" research in cognitive neuroscience. In reply, Gijs Holleman and his colleagues revived Hammond's critique (Holleman et al., 2020).

The alleged misuse of the concept of ecological validity had its origins in a classic paper by Martin Orne (1962) on the social psychology of research in experimental psychology. As a former graduate student of Orne's, I had long contemplated writing a response to Hammond's critique, so I decided to join the current thread with a note attempting to explicate what Orne had in mind (Kihlstrom, 2020). In it, I agreed that the common usage of "ecological validity", referring to something like mundane realism, is indeed inconsistent with Brunswik's intentions (and Orne's for that matter). But I also argued that Orne's revisionist usage is consistent with Brunswik's, because Orne emphasized the information provided to subjects about the true nature of an experiment – cues that Orne called "demand characteristics". An unedited preprint is available at https://www.ocf.berkeley.edu/~jfkihlstrom/PDFs/2020s/2020/EcologicalValidity_PPS_Rev1a_ref.pdf.

I do not know whether Orne ever met Brunswik. Orne, then based at Harvard Medical School, did teach briefly at Berkeley, but that was during the summer of 1962, after Brunswik's death (1955). On the other hand, Orne's family emigrated from Austria to New York at the time of the *Anschluss*, when he was about 11 years old. Orne's mother, Martha Brunner-Orne, a psychiatrist, may have known Brunswik and his wife in Vienna (Else Frenkel-Brunswik, a psychoanalyst who worked with Leon Festinger on *The Authoritarian Personality*, died in 1958).

Orne never offered his own definition of ecological validity, and his only reference citation to Brunswik (Orne, 1962, fn. 4) was to a long article that the latter had prepared for his course on research methods (it is difficult to imagine most of today's graduate students, much less undergraduates, getting through it). That document is still available in the UC Berkeley Library, and as far as I can determine it is identical to Brunswik's published contribution to a symposium on probability and statistics (Brunswik, 1947/1949). In his footnote, Orne characterizes ecological validity "in the sense that Brunswik (1947) has used the term: appropriate generalization from the laboratory to

nonexperimental situations". Such a phrase does not actually appear in the 1947 article, and the closest that Brunswik comes to it is in a discussion of *ecological* (or *situational*) *generality*, achieved through the *representative design* of psychological experiments in such a way as to insure that cues available in the real world are adequately sampled in the experimental setting.

For Brunswik, however, ecological generality is just a matter of insuring that, in an experiment on size constancy for example, the sizes and distances involved are an adequate sample of those that would be encountered in the real world outside the laboratory. Orne's revisionist concept of ecological validity goes beyond representative design. Instead, he argued that experiments sometimes contain cues – demand characteristics – that simply *aren't present at all* in the real-world setting. These cues are ecologically valid in the experimental setting, in that they provide information to a subject about the true nature of the experiment. But because they are unique to the experimental situation, they are not ecologically valid with respect to the real-world setting that actually interests the researcher. To the extent that subjects utilize demand characteristics in the experimental situation which have no counterpart in the real world, their behavior in the experiment will not generalize to the real world.

In part, Hammond's (1998) critique was justified: the familiar equation of "ecological validity" with mundane realism is indeed inconsistent with Brunswik's coinage. But it is also not what Orne (1962) had in mind. For Orne, experiments do not lack ecological validity when they fail to use life like stimulus materials and tasks, or even when they fail to predict behavior in the real world. Experiments lack ecological validity when they provide cues to the subject that the experimental situation is not what it appears to be, or as presented by the experimenter. Because the ecological validity of an experiment depends on the ecological validity of the cues it provides to the subject, Orne's revisionist construal is broadly consistent with Brunswik's intentions.

References:

Brunswik, E. (1947/1949). Systematic and representative design of psychological experiments with results in physical and social perception. In J. Neyman (Ed.), *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* (pp. 143–202). Berkeley. CA.: University of California Press.

Hammond, K. R. (1998). *Ecological validity: Then and now.* Retrieven from www.brunswik.org/notes/essay2.html

Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020). The Reality of "Real-Life" Neuroscience: A Commentary on Shamay-Tsoory and Mendelsohn (2019). *Perspectives on Psychological Science*, 1–5. https://doi.org/10.1177%2F1745691620917354

Kihlstrom, J. F. (2020). Ecological validity and "ecological validity". *Perspectives on Psychological Science*, in press.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*(11), 776–783. https://doi.org/10.1037/h0043424

Shamay-Tsoory, S. G., & Mendelsohn, A. (2019). Real-Life neuroscience: An ecological approach to brain and behavior research. *Perspectives on Psychological Science*, *14*(5), 841–859. https://doi.org/10.1177/1745691619856350

Sensemaking Support System (S^3) for Manufacturing Process Improvement

Thomas B. Ladinig & Krishna S. Dhir
Széchenyi István University, Győr, Hungary

Gyula Vastag
Corvinus University of Budapest & National University of Public Service, Budapest, Hungary

Contact: t.ladinig@yahoo.com

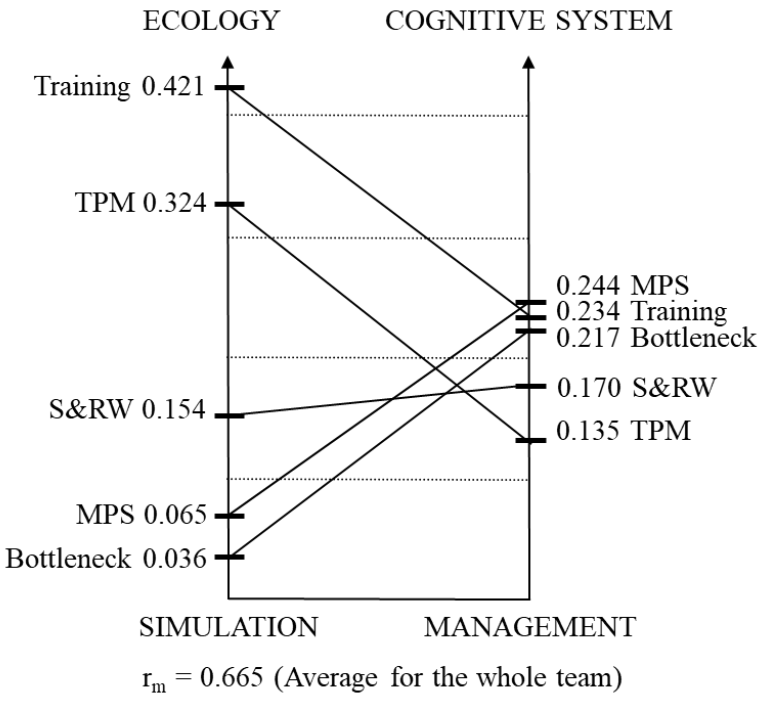
We apply social judgment theory and the lens model methodology for the development of a sensemaking support system to help operations managers find mutually acceptable solutions based on their individual judgments. This is done by comparing their judgments to each other and to the results of a simulation analysis conducted at a business unit of a multi-national car manufacturer.

The lens model has been used only scarcely in operations management and operations research (OM/OR) and we wanted to introduce this very interesting and useful methodology to our field once again. One very early publication from Ebert et al. (1985) can be found in the Journal of Operations Management, now one of the leading journals in our fields. The main inspiration for our paper was the work of Dhir (2001) where the usability and applicability of the lens model was showcased to improve judgments and decision making in OM/OR.

In our paper the lens model methodology was tied to the sensemaking process of Weick et al. (2005) to capture the behavioural aspects of managerial decision making and combine it with more traditional methods, like a discrete event simulation analysis. By making individual judgments visible, comparable, and understandable for the whole management team we helped managers to better understand their environment and their colleagues. We analysed individual judgments regarding five cues - typical processes within operations management - and their relative importance to improve the production system of the business unit in terms of manufacturing throughput time.

Our results indicate a lack of cohesion within the management team attempting to make mutually acceptable decisions, due to differences in judgments and individual preferences. Compared to the results of the simulation analysis (Figure 1) we found significant differences in judgments and lack of consensus within the management team to agree upon the most important cues. The lens model can help to uncover such differences and facilitate group decision making by reducing biases and errors in human decision making.

Figure 1. Comparison of weights between simulation and judgment analysis for five cues.



We conducted an a priori assessment and the participants could assign a total of 100 points to each of the five processes (cues) to represent their preferences before the actual judgment analysis. After that, we gave managers a total of 20 fictional production system configurations with different values for each cue which they had to rate on a scale from one to 20. Based on the results of the rating we conducted the judgment analysis and computed weights for each cue, the matching index, and the consistency of judgments for the twelve participants. Table 1 summarizes the results of the a priori assessment and the actual judgment analysis and uncovers that weights differ between a priori assessments and judgment analyses. This indicates biases and lack of understanding about own individual preferences in complex decision tasks. There is also a significant difference between weights for the same cue within the whole team. However, the average cue weight over all judges is less significant, compared to the simulation analysis, and shows a lack of consensus within the management team.

Table 1
Summary of results for the a priori assessment and judgment analysis

	A Priori Assessment					Judgment Analysis				
Name	**Training**	**TPM**	**Scrap&RW**	**MPS**	**Bottleneck**	**Training**	**TPM**	**Scrap&RW**	**MPS**	**Bottleneck**
TOPMGR1	0,3	0,2	0,1	0,2	0,2	0,37	0,07	0,10	0,15	0,31
NPIMGR1	0,15	0,23	0,15	0,23	0,24	0,16	0,19	0,18	0,36	0,12
NPIMGR2	0,3	0,1	0,3	0,2	0,1	0,21	0,10	0,15	0,30	0,24
PROMGR1	0,35	0,3	0,2	0,05	0,1	0,36	0,29	0,10	0,11	0,14
PROMGR2	0,2	0,2	0,2	0,15	0,25	0,16	0,14	0,23	0,26	0,22
PROMGR3	0,15	0,1	0,25	0,25	0,25	0,16	0,19	0,14	0,41	0,10
LOGMGR1	0,21	0,14	0,21	0,28	0,16	0,19	0,10	0,18	0,27	0,26
LOGMGR2	0,2	0,35	0,25	0,05	0,15	0,18	0,10	0,14	0,20	0,38
LOGMGR3	0,3	0,1	0,3	0,2	0,1	0,46	0,17	0,17	0,02	0,19
QUAMGR1	0,2	0,2	0,3	0,15	0,15	0,07	0,07	0,40	0,25	0,21
PRIMGR1	0,15	0,1	0,2	0,3	0,25	0,23	0,10	0,12	0,22	0,32
PRIMGR2	0,25	0,1	0,15	0,3	0,2	0,27	0,11	0,14	0,38	0,11
	Mean					Mean				
	0,230	0,177	0,218	0,197	0,179	0,235	0,135	0,170	0,244	0,217
	Std. Deviation					Std. Deviation				
	0,066	0,082	0,062	0,081	0,058	0,105	0,062	0,079	0,110	0,089
	Coefficient of Variation					Coefficient of Variation				
	0,286	0,464	0,286	0,414	0,324	0,447	0,462	0,464	0,451	0,409

The result of our work was a generally accepted action proposal for the top management of the business unit to improve specific processes based on the results of the judgment- and simulation analyses. Differences and similarities between judgments could be analysed and potential controversies resolved to define a common policy for all departments of the business unit. Future research could compare scientific interventions using the lens model methodology to purely quantitative analyses and study strength and weaknesses of both approaches. In general, we believe that scientific work in real-world industry settings could always benefit from including behavioural factors such as individual judgments and preferences.

For the full article, please see:
Ladinig, T. B., Dhir, K. S., & Vastag, G. (2020). Sensemaking support system (S3) for manufacturing process improvement. *International Journal of Production Research*, https://doi.org/10.1080/00207543.2020.1733700

References:

Dhir, K. S. (2001). Enhancing management's understanding of operational research models. *Journal of the Operational Research Society, 52*(8), 8735–887. https://doi.org/10.1057/palgrave.jors.2601148

Ebert, R. J., Rude, D. E., & Cecil, E. A. (1985). Capturing judgments to clarify production strategy and policy. *Journal of Operations Management 5*(2), 129–149. https://doi.org/10.1016/0272-6963(85)90003-8

Weick, K. E., Sutcliffe, K. M., &, Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organizational Science, 14*(4), 409–421. https://doi.org/10.1287/orsc.1050.0133

Clinical Judgments of Recidivism by Prison Staff and Official Recidivism Rates of Female Prisoners: An Application of Brunswik's Lens Model in a Women's Prison in Lower Saxony, Germany

Susann Prätor & Marcel Guéridon
Criminological Service Unit of Lower Saxony, Germany

Contact: Susann.Praetor@justiz.niedersachsen.de

This paper summarizes main results of our comprehensive article "Can prison staff predict recidivism of female prisoners? – An application of Brunswik's lens model in a women's prison in Lower Saxony, Germany" which was published in *Forensische Psychiatrie, Psychologie, Kriminologie* in 2020 (Prätor & Guéridon, 2020). Based on the assumptions of Brunswik's lens model (1955) and social judgment theory (Hammond et al., 1975) clinical judgments on prisoners' dangerousness by prison staff were compared to official recidivism of female prisoners. Main questions of this analysis were: 1) In which way do both measures correspond to each other? and 2) Are both measures predicted by the same variables?

The *judgments* of women's risk of recidivism by prison staff were based on a survey among prison staff on a total of 294 women who were detained in prison in Lower Saxony (Prätor & Suhling, 2016). Data on actual recidivism within three years after release (*criterion*) were received from the National central register. The survey among prison staff also provided data on independent variables (*cues*). These cues encompass sociodemographic and criminological factors as well as drug addiction (and its change during imprisonment) and circumstances after release, e.g., social support, work/job availability (for an overview c.f. Andrews et al., 2012; Bonta & Andrews, 2017; Greiner et al., 2015; Rodermond et al., 2015)

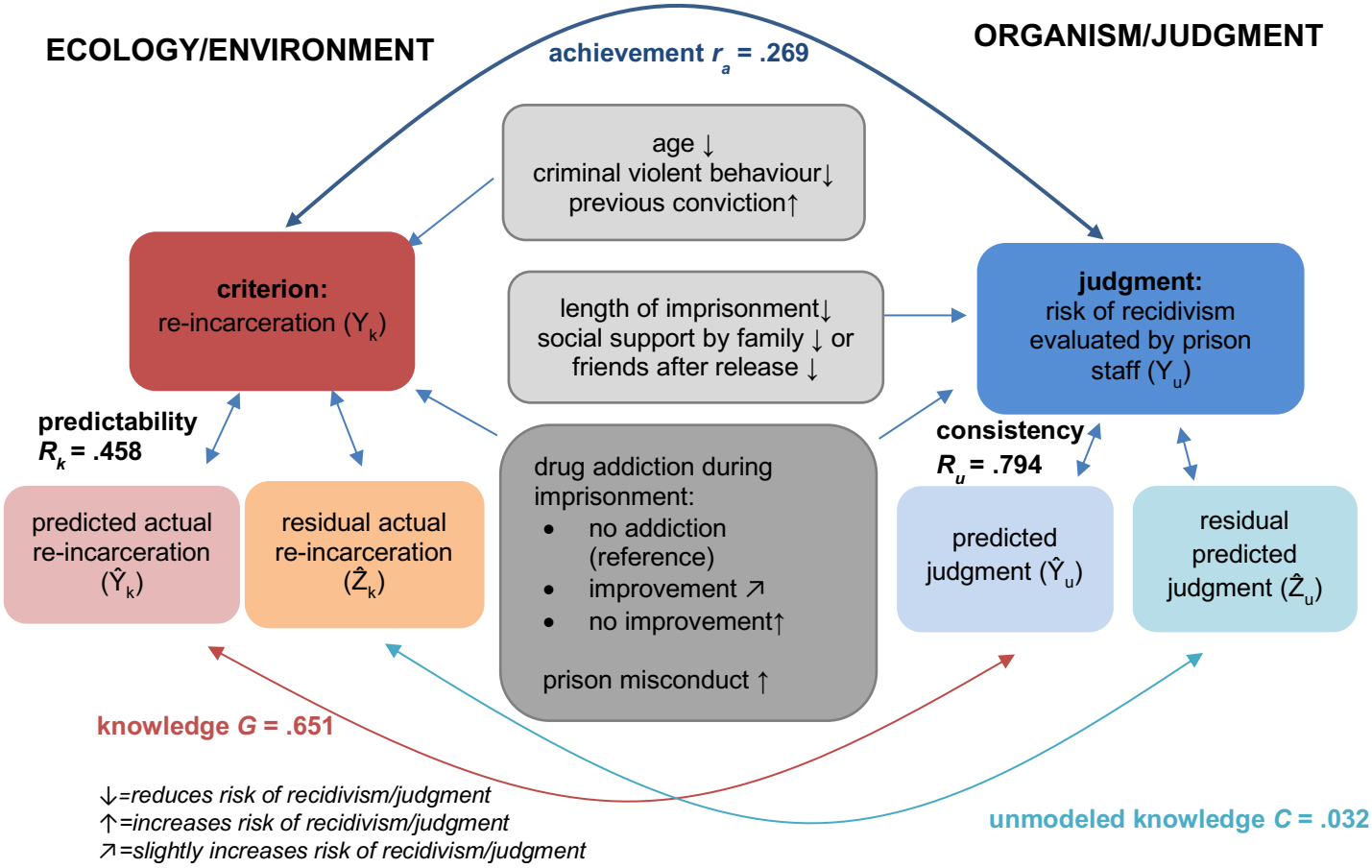
Firstly, we analysed the correlation between judgement of recidivism risk by prison staff and official recidivism (in terms of re-incarceration) of female prisoners three years after release. As expected a higher risk of recidivism attested by prison staff is connected to the risk of official recidivism (Cramers V = .306, p < .001, N = 276).[1] Judgment was measured by asking prison staff to estimate a women's individual risk of recidivism on a 5-point scale (no risk, rather low, some risk, rather high, high).[2] In order to combine official recidivism rates with judgments this variable was dichotomized into "no risk/rather low risk" vs. "some risk to high risk". In most cases (62.7%), judgment and objective criterion coincide with each other (39.1% negative and 23.6% correct positive). With regard to

[1] We only refer to re-incarceration (instead of reconviction) in this paper because of its stronger correlation with judgment by prison staff.
[2] Relative frequencies of this item are: 20.3% low risk, 26.1% rather low risk, 30.8% some risk, 17.8% rather high risk, 5.1% high risk).

incorrect judgments the rate of false positive results is more than four times higher than the rate of false negative results (30.1% vs. 7.2%).

According to Brunswik's lens model and social judgment theory we then calculated different models (Figure 1).

Figure 1. Prediction of re-incarceration (actual recidivism and judgement by prison staff) according to Brunswik's lens model.



On the one hand, there is a model of the ecology/environment using official data of re-incarceration by female prisoners within three years after release predicted by different cues (logistic regression). On the other hand, a model for judgment was calculated using identical variables (linear regression).[3]cResults show that some variables (e.g., age or previous convictions) only predict actual re-incarceration while length of imprisonment and social support by family or friends after release are only significant predictors for judgment. Drug addiction and its improvement during imprisonment and prison misconduct play an important role for judgment by prison staff as well as for the objective criterion. Female prisoners with no improvement in the domain of drug addiction during

[3]This model is not a classical lens model with two multiple regressions, but a hybrid model because of combination of linear with logistic regression parameters (Hamm & Yang, 2017).

imprisonment have the highest risk of re-incarceration in comparison to those prisoners without drug addiction problems. There is also a higher risk of recidivism for female prisoners with only slight improvements. Another cue that is predictive for both indicators is prison misconduct. Prisoners who were labelled by prison staff as violating prison rules have a higher risk of reoffending in the eyes of prison staff and according to official recidivism rates.

In Figure 1 the central components of the lens model as usually used in adaptation to social judgement tasks are reported (achievement, predictability, knowledge, consistency and unmodeled knowledge) and can be compared with other studies (Karelaia & Hogarth, 2008; Kaufmann et al., 2013). The *predictability* of R_k = .458 shows that accurate prediction of re-incarceration by these cues is rather limited – also in comparison to other lens model studies. Therefore, the amount of *achievement* (r_a = .269) as the correlation between criterion and subjective judgment given the predictors is also limited – even if prison staff would use all of these indicators in a highly consistent manner. This value is in line with results of Kaufmann et al. (2013) but low in comparison to values reported in (non-)psychological studies using Brunswik's lens model (Karelaia & Hogarth, 2008). However the independent variables used in the judgment model are appropriate for predicting judgment of recidivism and are consistently used by prison staff (*consistency* R_u = .794*)*.[4] In comparison with other psychological studies, *knowledge* as congruence of ecology/environment and judgment model (*G* = .651) is high (Kaufmann et al., 2013). However, in a meta-analysis of lens model studies in different (non)psychological areas by Karelaia and Hogarth (2008), higher figures for knowledge could be found. In spite of different limitations of this study (e.g., measurement of subjective judgment, only official data on recidivism, limited cues), Brunwik's lens model as well as social judgment theory proved again to be promising theoretical models in order to analyse decision-making processes in the field of criminal justice.

For the full article, please see:

Prätor, S., & Guéridon, M. (2020). Wie gut können Justizvollzugsbedienstete das Rückfallrisiko von Inhaftierten vorhersagen? [How well can prison staff predict recidivism of female prisoners? An application of Brunswik's lens model in a woman's prison in Lower Syxony, Germany]. *Forensische Psychiatrie, Psychologie, Kriminologie, 14*(3), 315–327. https://doi.org/10.1007/s11757-020-00608-x

References:

Andrews, D. A., Guzzo, L., Raynor, P., Rowe, R. C., Rettinger, L. J., Brews, A., & Wormith, J. S. (2012). Are the major risk/need factors predictive of both female and male reoffending? A test with the eight domains of the level of service/case management inventory. *International Journal of Offender Therapy and Comparative Criminology, 56*(1), 113–133. https://doi.org/10.1177/0306624X10395716

Bonta, J., & Andrews, D. A. (2017). *The Psychology of Criminal Conduct*. London/New York: Routledge.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*(3), 193–217. https://doi.org/10.1037/h0047470

Greiner, L. E., Law, M. A., & Brown, S. L. (2015). Using dynamic factors to predict recidivism among women. A four-wave prospective study. *Criminal Justice and Behavior, 42*(5), 457–480. https://doi.org/10.1177/0093854814553222

[4] As we could not connect judgments to specific judges, this estimation as well as the estimated achievement is probably too high compared to more appropriate analyses on the level of single judges as requested by Brunswikian principles.

Hamm, R. M., & Yang, H. (2017). Alternative lens model equations for dichotomous judgments about dichotomous criteria. *Journal of Behavioral Decision Making, 30*(2), 527–532. https://doi.org/10.1002/bdm.1969

Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social Judgment Theory. In M. F. Kaplan & S. Schwartz (Eds.), *Human Judgment and Decision Processes.* New York: Academic Press.

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin, 134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Kaufmann, E., Reips, U.-D., & Wittmann, W. W. (2013). A critical meta-analysis of lens model studies in human judgment and decision-making. *PLoS ONE, 8*(12), 1–16. https://doi.org/10.1371/journal.pone.0083528

Prätor, S., & Guéridon, M. (2020). Wie gut können Justizvollzugsbedienstete das Rückfallrisiko von Inhaftierten vorhersagen? *Forensische Psychiatrie, Psychologie, Kriminologie, 14*(3), 315–327. https://doi.org/10.1007/s11757-020-00608-x

Prätor, S., & Suhling, S. (2016). Legalbewährung von Frauen - Befunde einer Untersuchung im niedersächsischen Frauenvollzug. *Monatsschrift für Kriminologie und Strafrechtsreform, 99*(3), 215–236.

Rodermond, E., Kruttschnitt, C., Slotboom, A.-M., & Bijleveld, C. (2015). Female desistance: A review of the literature. *European Journal of Criminology*, *13*(1), 3–28. https://doi.org/10.1177/1477370815597251

A Brunswikian View on Judging Fake News in Social Media Communications

---

Christian Unkelbach
University of Cologne, Germany

Adrien Mierop & Olivier Corneille
Catholic University of Louvain, Belgium

Contact: christian.unkelbach@uni-koeln.de

How people judge the truth of information has resurfaced as a central question for theoretical and applied psychology (see Brashier & Marsh, 2020). This re-emergence is fueled by the question how and why people may believe "fake news" on social media (Vosoughi et al., 2018). We argue that such truth judgments are best understood as a case of Brunswik's (1952) lens model: truth by itself is inaccessible, it is a *distal* criterion. Thus, the organism must infer truth based on *proximal*, or accessible cues (see Karelaia & Hogarth, 2008). The question of interest for empirical psychologists is then which cues people use for this task (i.e., people's cue utilization), and not so much which cues factually predict truth (i.e., the cue's ecological validity; see Unkelbach et al., 2019).

One prominent cue for people to judge truth is processing fluency (Reber & Schwarz, 1999; Unkelbach & Rom, 2017). Typically, people believe information more when they process it more fluently. Unkelbach and Greifeneder (2013, 2018) presented a variant of the lens model that posits that people use fluency cues in accordance with their ecological validity (see Reber & Unkelbach, 2010). Because truth is generally more prevalent than falsehood, this model predicts that people adaptatively use fluency as a cue for truth in standard judgment contexts. Of importance, however, this model allows predicting reversals of the fluency influence in specific judgment ecologies, such as when people

judge fluently processed information as “false” instead of “true” (Unkelbach, 2007), fluently generated information as “negative” instead of “positive” (Briñol et al., 2006), or fluently processed names as “rare” instead of “frequent” (Oppenheimer, 2004).

However, there are also other models that explain why people generally believe fluently processed information more. First, processing information fluently feels good or positive, which is referred to as the *hedonic marking hypothesis* (Winkielman et al., 2003). And because truth is also positive, people might infer that fluently processed information is true (Unkelbach et al., 2011). Second, fluency might amplify any judgment that is presented to participants, which is referred to as the *amplification hypothesis* (Albrecht & Carbon, 2014).

Corneille et al. (2020) contrasted predictions for truth judgments from the *lens model* with predictions from the *amplification* and the *hedonic marking* hypothesis (see Table 1). They compared the effects of repetition-induced fluency in standard and social media ecologies. In standard judgment contexts, the models converge (Table 1’ s column 1): As just noted, because truth is prevalent in standard ecologies, the lens model predicts that repeating information generally enhances its perceived truth. Likewise, because processing information fluently feels good or positive, and because truth is also positive, people might infer that fluently processed information is true. Finally, because fluency might enhance any judgment, people should judge fluently processed information as “true” when asked for truth judgments.

The models diverge when considering falsehood judgments (Table 1’ s column 2). Typically, participants judge if a statement is true or false. If one modifies this to participants judging if a statement is false or not, the amplification hypothesis predicts that repetition-induced fluency increases judgments of *falsehood*, because repetition should amplify any judgment. The lens model and the hedonic marking hypothesis predict no or a weakening effect of repetition, because what is true cannot be false, for the lens model; and because falsehood is negative, for the hedonic marking hypothesis.

The most interesting case is a switch in the ecology, when participants judge prior use of the statements as fake news on social media (Table 1’ s column 3). If people indeed adaptively learn that fluency might indicate falseness in social media communications, then they might use fluency as a cue to judge information as fake news in the ecology of social media. Thus, the lens model predicts that repetition-induced fluency increases judgments of previous use of the statements as fake news on social media. The same is true for the amplification hypothesis, which again, predicts increases for any judgments. However, the hedonic marking predicts decreases or no effect, as “fake news” is obviously a negative response. As Table 1 illustrates, across the conditions, each model has a unique pattern of predictions.

Across three experiments, Corneille et al. (2020) found clear support for an ecological lens model: when participants judged whether repeated information is “true”, they believed repeated information more, as the rate of “true” judgments increased (Column 1 in Table 1). When participants judged whether repeated information is “false”, participants still believed repeated information more, as the rate of “false” judgments *decreased* (Column 2 in Table 1). Critically, when participants judged if information was used as fake news in social media communications, the rate of “fake news” ratings *increased* (Column

3 in Table 1).

Table 1
Differential predictions from the hedonic marking hypothesis, the amplification hypo-thesis, and a Brunswikian lens model for the effect of fluency on judgments of truth, falsehood, and use as fake news on social media

Context	Judgments "True" (Unspecified)	"False" (Unspecified)	"Fake News" (Social Media)
Hedonic marking	Increases	No effect/Reduces	No effect/Reduces
Amplification	Increases	Increases	Increases
Lens Model	Increases	No effect/Reduces	Increases

Note. The "Truth" and "False" judgments refer to standard ecologies.

These results provide evidence for an ecological lens model (Brunswik, 1952). They also provide a positive outlook on people's adaptive behavior. One may speculate based on these data that people indeed learn that repetition and the resulting fluency signal fakeness rather than truth in a social media ecology. Thus, instead of uncritically believing what they hear repeatedly, people may learn the ecological validity of a given cue (here: processing fluency) and adjust their cue utilization accordingly, which may lead to people judging repeated information as "fake news".

References:

Albrecht, S., & Carbon, C.- C. (2014). The Fluency Amplification Model: Fluent stimuli show more intense but not evidently more positive evaluations. *Acta Psychologica*, *148*, 195–203. https://doi.org/10.1016/j.actpsy.2014.02.002

Brashier, N. M., & Marsh, E. J. (2020). Judging truth. *Annual Review of Psychology*, *71*, 499–515. https://doi.org/10.1146/annurev-psych-010419-050807

Briñol, P., Petty, R. E., & Tormala, Z. L. (2006). The malleable meaning of subjective ease. *Psychological Science*, *17*(3), 200–206. https://doi.org/10.1111/j.1467-9280.2006.01686.x

Brunswik, E. (1952). The conceptual framework of psychology. *International Encyclopedia of Unified Science*, *1*(10), University of Chicago Press.

Corneille, O., Mierop, A., & Unkelbach, C. (2020). Repetition increases both the perceived truth and fakeness of information: An ecological account. *Cognition*, *205,* 1–6. https://doi.org/10.1016/j.cognition.2020.104470

Karelaia, N., & Hogarth, R. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, *134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Oppenheimer, D. M. (2004). Spontaneous discounting of availability in frequency judgment tasks. *Psychological Science*, *15*(2), 100–105. https://doi.org/10.1111/j.0963-7214.2004.01502005.x

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*(3), 338–342. https://doi.org/10.1006/ccog.1999.0386

Reber, R., & Unkelbach, C. (2010). The epistemic status of processing fluency as source for judgments of truth. *Review of Philosophy and Psychology*, *1*, 563–581. https://doi.org/10.1007/s13164-010-0039-7

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 219–230. https://doi.org/10.1037/0278-7393.33.1.219

Unkelbach, C., & Greifeneder, R. (2013). A general model of fluency effects in judgment and decision making. In C. Unkelbach & R. Greifeneder (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behaviour* (pp. 11–32). New York, NY: Psychology Press.

Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, *79*, 78–86. https://doi.org/10.1016/j.jesp.2018.06.010

Unkelbach, C., & Rom, S. C. (2017). A referential theory of the repetition-induced truth effect. *Cognition*, *160*, 110–126. https://doi.org/10.1016/j.cognition.2016.12.016

Unkelbach, C., Bayer, M., Alves, H., Koch, A., & Stahl, C. (2011). Fluency and positivity as possible causes of the truth effect. *Consciousness and Cognition*, *20*(3), 594–602. https://doi.org/10.1016/j.concog.2010.09.015

Unkelbach, C., Koch, A., Silva, R. R., & Garcia-Marques, T. (2019). Truth by repetition: Explanations and implications. *Current Directions in Psychological Science*, *28*(3), 247–253. https://doi.org/10.1177/0963721419827854

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Mahwah, NJ: Lawrence Erlbaum.

Pushing the Limits for Judgmental Consistency: Comparing Random Weighting Schemes with Expert Judgments

Martin C. Yu
HumRRO, USA

Nathan R. Kuncel
University of Minnesota, USA

Contact: myu@humrro.org

We study the degree to which inconsistency in combining information when making multiple judgments is detrimental to the predictive validity of expert judgment. Because judgmental processes involve two aspects of data combination – the optimality of the data combination policy and the consistency with which the policy is applied – it would be necessary to tease apart consistency from optimality if the effects of consistency are to be studied. This can be done by examining random weighting schemes as there is no expectation of optimality, and pitting expert judgment against random weights in combining predictor information. When the intent is to make the most accurate judgment possible, randomly weighting information cues to make a judgment is the opposite of using a set of optimal regression weights.

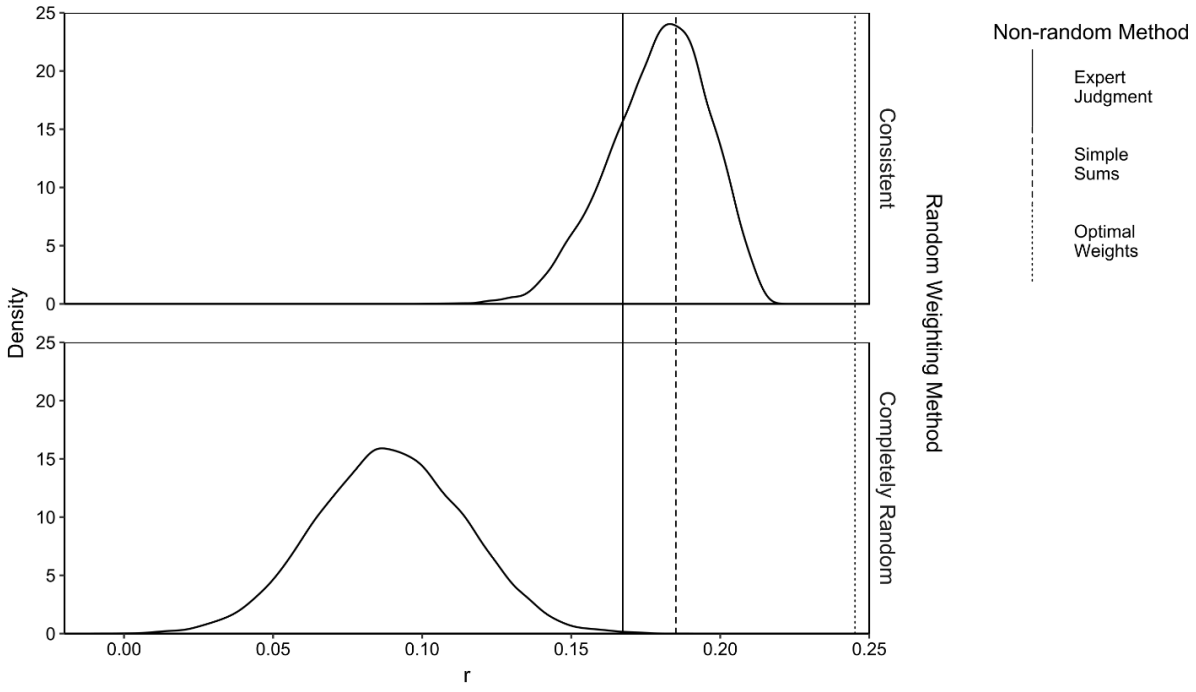
There are two forms of random weighting that warrant consideration. The first form was used by Dawes and Corrigan (1974), where a set of random weights is generated, and applied consistently to every single judgmental case. In a simulation study, this is repeated many times so that the average validity of consistent use of random weights can be estimated. The second form is completely random weighting, where a set of random weights is generated for every single judgmental case. Here, no two judgments are combined using the same weighting policy (unless by coincidence). Again, this process is repeated many times to estimate the average validity of truly random weighting. With consistent random weights, there is no expectation of optimality, but there is an expectation of consistency. With completely random weights on the other hand, there is no expectation of either optimality or consistency.

In this study, these two random weighting approaches are applied through a Monte Carlo simulation. In contrast to Dawes and Corrigan (1974), where only the average

validity of consistent random weights was evaluated, this study additionally examines the average validity of completely random weights, as well as the complete distributions of validities across all simulation trials for these two random weighting approaches. By comparing the validity of subjective expert judgment and non-random mechanical methods such as unit weighting via simple sums and optimal regression weights in the context of the distributions of random but consistent and truly random weighting, we can more precisely determine the extent to which non-random methods of prediction outperform or do not outperform these random methods.

We conducted our analyses on three separate assessment validation datasets where job candidates were evaluated for management positions by doctoral-level psychologists trained in conducting managerial hiring assessments. Weights were constrained to be positive as all assessment predictors were expected to be positively related to potential job performance. Overall conclusions were similar across all three datasets, and results for one dataset, Company A, is presented here (see Figure 1). Complete results can be found in Yu and Kuncel (2020).

Figure 1. Density distributions of validities (10,000 iterations each) at Company A of predictor scores combined using random positive weights applied consistently across all candidates (top plot) and completely random positive weights (bottom plot) generated for each candidate. Vertical lines are validities at Company A of non-random methods of data combination: expert judgment (solid line), unit weighting via simple sums (dashed line), and optimal weighting (dotted line).



When the overall ratings computed using random methods were used to predict job performance at Company A, across 10,000 iterations, random weights applied consistently across candidates had a mean predictive validity of r = .18 (SD = .02), and ranged from r = .10 to .22. Random weights applied consistently outperformed expert judgments in 76.83% of the iterations, simple sums in 39.40% of the iterations, and never outperformed optimal weights. Completely random weighting across candidates had a mean validity of r = .09 (SD = .02), and ranged from r = -.01 to .19. Completely random

weights never outperformed expert judgment, simple sums, or optimal weights. 69.85% of the iterations for completely random weights were outperformed by all of the iterations for random weights applied consistently.

These results indicate that experts do not make judgments completely randomly and are aware, to some extent, of what information is most valuable. However, their inconsistency in combining information does drastically damage their accuracy. This simulation study demonstrates that consistency in applying predictor weights is paramount to making accurate judgments. It is striking that mindless consistency is enough to result in more accuracy than expert judgment. On average, random weights applied consistently resulted in better predictions than the assessors' own judgments, which parallels Dawes and Corrigan's (1974) earlier study of random weighting.

Ultimately, the finding that even random weights perform well when applied consistently suggests that consistency in applying predictor weights is more important than the weights themselves. Linear models are quite robust, and as long as the signs on the weights do not change, changes in weights are not expected to drastically impact their predictive power (Dawes, 1979). As Waller (2008) demonstrated with fungible weights, it is possible to derive an infinite number of alternate regression weighting schemes that yield a predictive validity almost as good as that of optimal weights (in multiple regression with three or more predictors). That being said, even though it is possible to generate a set of random weights that will perform very well when applied consistently, it can be difficult or impossible to tell how well that set of random weights will perform until the validation is conducted. In this simulation study, both optimal and unit weights via simple sums tend to perform better than random weights applied consistently. Practically speaking, if optimal weights are not known or cannot be approximated, it would be better to simply add up predictor scores instead of using an ill-defined weighting scheme.

For the full article, please see:

Yu, M. C., & Kuncel, N. R. (2020). Pushing the limits for judgmental consistency: Comparing random weighting schemes with expert judgments. *Personnel Assessment and Decisions, 6*(2), 1–10. https://doi.org/10.25035/pad.2020.02.002

References:

Dawes, R., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*(2), 95–106. https://doi.org/10.1037/h0037613

Dawes, R. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34*(7*),* 571–582. https://doi.org/10.1037/0003-066X.34.7.571

Waller, N. G. (2008). Fungible weights in multiple regression. *Psychometrika, 73*(4), 691–703. https://doi.org/10.1007/s11336-008-9066-z

Yu, M. C., & Kuncel, N. R. (2020). Pushing the limits for judgmental consistency: Comparing random weighting schemes with expert judgments. *Personnel Assessment and Decisions, 6*(2), 1–10. https://doi.org/10.25035/pad.2020.02.002

Improving Judgment Accuracy by Sequential Adjustment

Shenghua Luan
Institute of Psychology, Chinese Academy of Sciences

Lael J. Schooler
Department of Psychology, Syracuse University

Jolene H. Tan
Max Planck Research Group Naturalistic Social Cognition, Max Planck Institute for Human Development

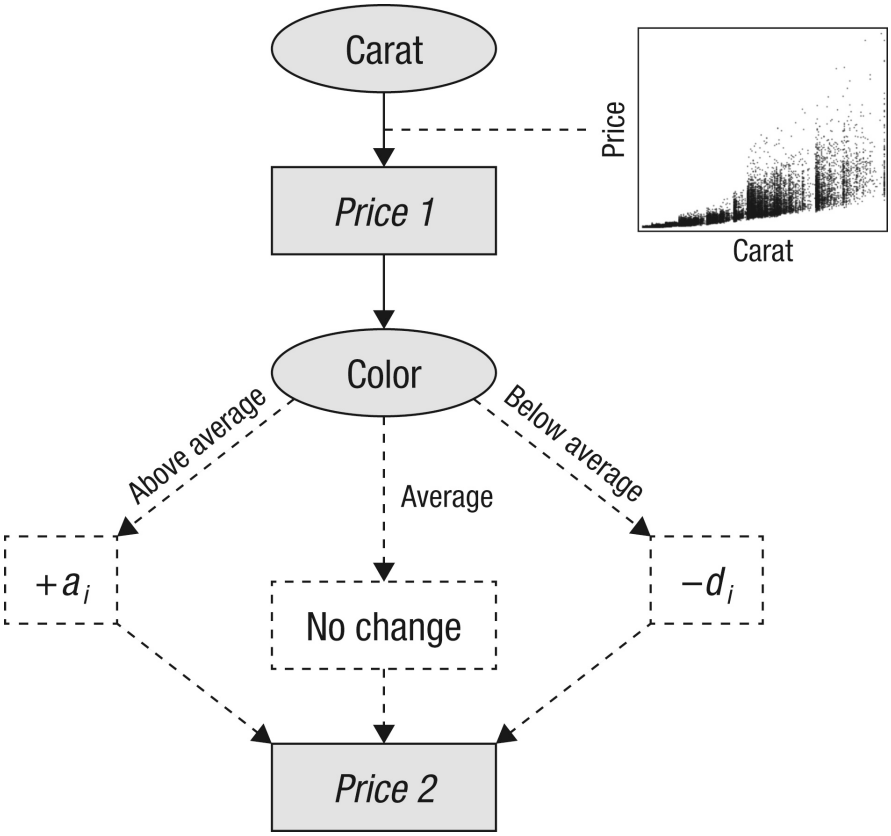
Contact: luansh@psych.ac.cn

Previous studies show that our judgment accuracy is often well below the ideal level, and the main problem seems to be the difficulty of integrating information from multiple cues (e.g., Karelaia & Hogarth, 2008). In this study (Luan, Schooler, & Tan, 2020), we proposed a simple method that may improve judgment accuracy.

To assess judgment accuracy, researchers typically present all cues at once and ask participants to provide an estimate afterward. This "simultaneous" procedure, in terms of cue presentation, promotes a weighting-and-adding approach to information integration, because of people's strong tendency to weight-and-add in the absence of the need to search for and update information (e.g., Bröder, 2011; Rieskamp & Otto, 2006). Meanwhile, in judgment studies where accuracy is difficult to assess (e.g., personal impressions), a "sequential" procedure, in which cues are presented one by one and participants are asked to provide an estimate at each step, has also been applied. This procedure promotes a "sequential adjustment" approach to information integration (e.g., Hogarth & Einhorn, 1992; Juslin et al., 2008). An example of this approach applied to judging diamond price is illustrated in Figure 1.

Suppose that a person applies the approach to estimate the price of a 0.61-carat diamond with an above-average color grade of "F." The person might first give a ballpark estimate of $3,500 based on the carat and then add another $300 after the color grade is revealed. Errors will occur if the carat-to-price function is not an accurate one and an inappropriate adjustment value is applied. However, in comparison to weighting-and-adding, this approach can reduce the difficulty of cue integration in several aspects: (a) All operations are performed on variables' original values; (b) there is no need to align drastically different cues by assigning them weights that are difficult to interpret psychologically and to learn; and (c) relatively little computation is needed to carry out the process, which can lower operational errors that harm judgment accuracy.

Figure 1. An illustration of the sequential adjustment process using the example of estimating diamond price based on first the carat and then the color cue. The graph in the figure depicts the relationship between carat and price in the diamond data set collected in Study 2.



We hypothesized that people may judge more accurately by following a sequential than a simultaneous procedure, because the sequential procedure can reduce the difficulty of cue integration. To test this hypothesis, we conducted four studies in which we asked both experienced and inexperienced participants for their judgments in two task domains.

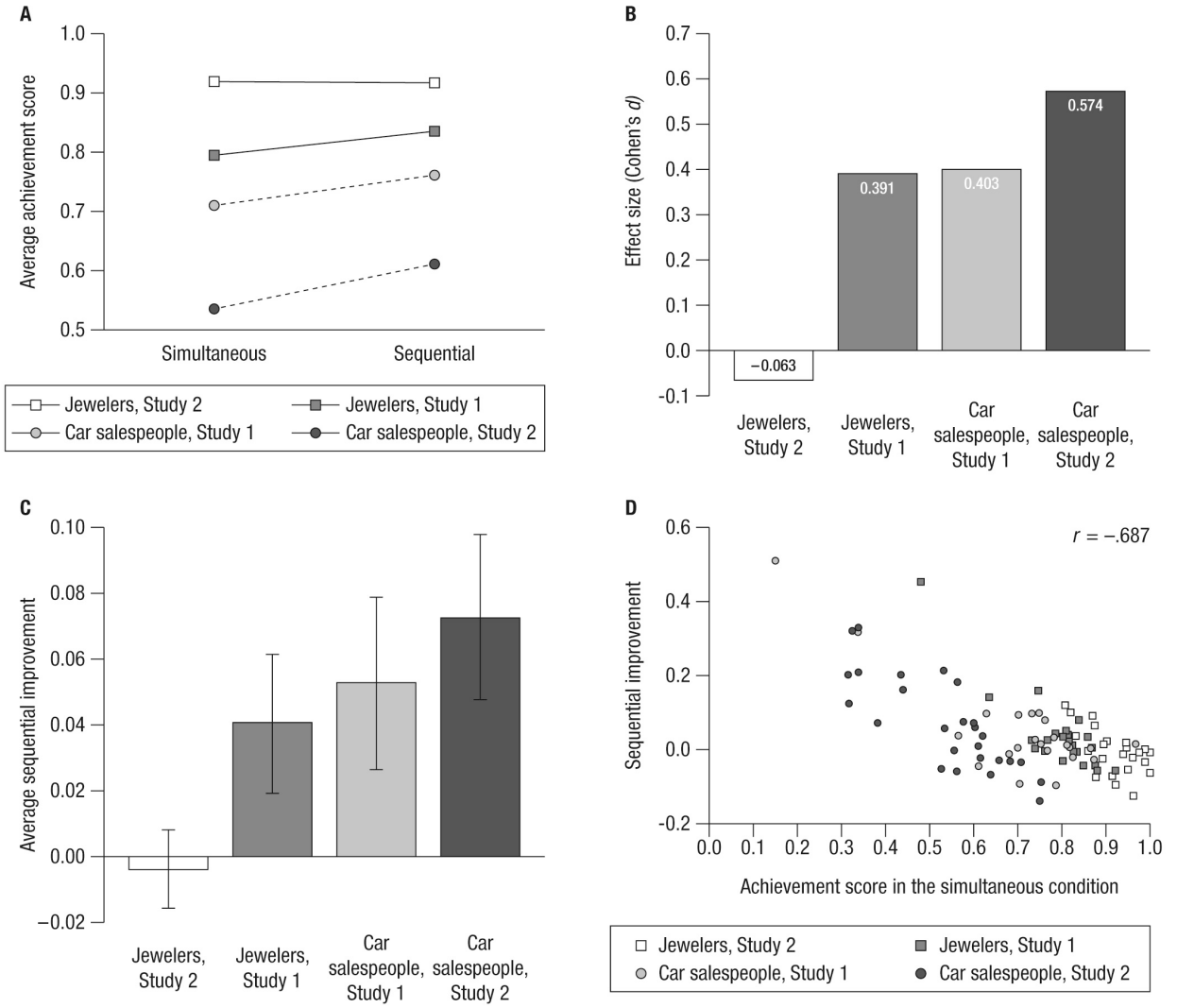
Participants in Studies 1 and 2 were working professionals. The studies had the same design but were conducted in different years. In each study, we asked jewelers to judge the price of diamonds on the basis of carat and color, and car salespeople to judge the fuel economy of cars on the basis of horsepower and number of cylinders. The two tasks differed much in linear predictability, each cue's ecological correlation, and the inter-cue correlation. The experimental materials were representative samples of real-world data collected in the year of the study.

A within-subjects design with two conditions was applied in each task. In the simultaneous condition, participants were given values of two cues simultaneously and asked to provide their estimates of the criterion variable afterward. In the sequential condition, participants were first shown the value of one cue—carat for the diamond task and horsepower for the car task—and asked to give an initial estimate of the criterion variable; after that, they were given the second cue's value and were asked to make a

second, final estimate. Participants' payoff depended on the accuracy of their judgments, and feedback was provided after each trial.

Figure 2 shows the key results of these two studies. In general, participants judged more accurately in the sequential condition than in the simultaneous condition, except for the jewelers in Study 2, whose judgment accuracy was exceptionally high in both conditions. Across participants, there was a negative correlation between accuracy in the simultaneous condition and the "sequential improvement" (i.e., the achievement score in the sequential condition minus the score in the simultaneous condition) of a participant. If we assume that performance in the simultaneous condition is the "default" performance, this implies that the lower one's default performance, the more the person stands to gain by judging with the sequential procedure.

Figure 2. Judgments by the professionals. (A) The average achievement score of each participant group in each experimental condition. (B) The effect size, in Cohen's d, of the experimental manipulation in each participant group, ordered by magnitude from left to right. (C) The average sequential improvement, that is, the difference between the sequential and simultaneous conditions (former minus latter) in a participant's achievement scores, for each group. Error bars indicate ± 1 SE. (D) A scatterplot of all participants' achievement scores in the simultaneous condition and their sequential improvements (N = 98). Shown also is the Pearson correlation, r, between the two variables across all participants.



In Studies 3 and 4, we conducted a similar experiment with college student participants and a between-subjects design. The results confirm the finding that judging in the sequential condition led to higher accuracy than in the simultaneous condition for both the diamond and the car tasks. Moreover, we added a "sequential optional" condition in Study 3, in which participants could choose which cue to check first. The accuracy in this condition was not different from the "sequential-fixed" condition and was better than the simultaneous condition. Finally, we conducted lens model analyses of participants' judgments in all studies. The results show that the sequential improvement effect was mainly caused by improved consistency scores, consistent with our analysis and hypothesis.

In sum, in three studies that covered two task domains and involved both experienced and inexperienced participants, we demonstrated how a simple twist in the judgment procedure could improve judgment accuracy. Participants were given the same cue information in the simultaneous and sequential conditions; yet, most in the sequential condition achieved higher accuracy. The method was less effective for participants who could judge more accurately in the simultaneous condition, but they were also the ones who had less room or need to improve.

For the full article, please see:

Luan, S., Schooler, L. J., & Tan, J. H. (2020). Improving judgment accuracy by sequential adjustment. *Psychonomic Bulletin & Review, 27*, 170–177. https://doi.org/10.3758/s13423-019-01696-5

References:

Bröder, A. (2011). The quest for take-the-best: Insights and outlooks from experimental research. In G. Gigerenzer, R. Hertwig, & T. Pachur (Eds.), *Heuristics: The foundations of adaptive behavior* (pp. 364–382). New York, NY: Oxford University Press.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24*, 1–55. https://doi.org/10.1016/0010-0285(92)90002-J

Luan, S., Schooler, L. J., & Tan, J. H. (2020). Improving judgment accuracy by sequential adjustment. *Psychonomic Bulletin & Review, 27*, 170–177. https://doi.org/10.3758/s13423-019-01696-5

Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple cue judgment: A division of labor hypothesis. *Cognition, 106*, 259–298. https://doi.org/10.1016/j.cognition.2007.02.003

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin, 134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*(2), 207–236. https://doi.org/10.1037/0096-3445.135.2.207

The Brunswik Society

26th International (Virtual) Meeting of the Brunswik Society
3-4 December 2020, 12.00-13.30 EST (17.00-18.30 GMT) via Zoom

3rd December, 12.00-13.30 EST (17.00-18.30 GMT) via Zoom
Expert Judgment and the Lens Model

Title: Hybrid Lens Model Analysis of Judgments by Medical Society Expert Guideline Panels: How Does Group Discussion Change Individuals' Judgments? **Authors:** Robert M. Hamm, Marianne Razavi, Iztok Hozo, Gordon H. Guyatt, Benjamin Djulbegovic, University of Oklahoma Health Sciences Center

Title: A Brunswikian Theory of Expert Judgment and its Application to Probabilistic Forecasting. **Authors:** Fergus Bolger, Anglia Ruskin University

Title: Choose your words wisely: semiotic characteristics predict halo effects and attrition in the serial reproduction of person descriptions. **Authors:** Karolin Salmen and Klaus Fiedler, Heidelberg University

Title: Judgements and Decision-Making on prisoners' dangerousness - An SJT Perspective. **Authors:** Marcel Guéridon & Susann Prätor, Criminological Service Unit of Lower-Saxony, Germany

4th December, 12.00-13.30 EST (17.00-18.30 GMT) via Zoom
Research Design and Adaptive Cognition

Title: Ecological validity and the 'real world' in the field of social attention. **Authors:** Gijs A. Holleman, Utrecht University

Title: Representative Design in Psychological Assessment: A Case Study Using the Balloon Analogue Risk Task (BART). **Authors:** Markus D. Steiner and Renato Frey, University of Basel

Title: The ecology of competition: A theory of risk--reward environments in adaptive decision making. **Authors:** Timothy J. Pleskac (University of Kansas), Larissa Conradt (Max Planck Institute for Human Development), Christina Leuker (Max Planck Institute for Human Development), and Ralph Hertwig (Max Planck Institute for Human Development)

Title: Precise/Not Precise (PNP): A Brunswikian Model that Uses Judgment Error Distributions to Identify Cognitive Processes. **Authors:** Joakim Sundh, August Collsiöö, Philip Millroth & Peter Juslin, Uppsala University

Social Hour!! via Interactive Online Platform Simulating the Real Experience
4th December 2020, 13.45 EST (18.45 GMT)
Free event – invites will be sent out to all meeting delegates

Hybrid Lens Model Analysis of Judgments by Medical Society Expert Guideline Panels: How Does Group Discussion Change Individuals' Judgments?

Robert M. Hamm, Marianne Razavi, Iztok Hozo, Gordon H. Guyatt, & Benjamin Djulbegovic
Oklahoma Health Sciences Center, USA

Contact: Robert-Hamm@ouhsc.edu

We applied the Hybrid Lens Model to analyze data from participants in expert panels convened by medical specialty societies for the review of proposed guidelines for clinical practice. Before and after panel meeting, each panel member was asked to evaluate each proposal, rate it on four features (evidence certainty, value certainty, benefit-harm balance, and costs) and state their recommendation on a five-point scale (strongly recommend to strongly oppose). During the panel meeting, they made such judgments again but we have only the consensus recommendation for each proposal. Post-meeting recommendations agreed more than pre with the meeting consensus, and cue judgments changed. We compare guideline recommendations before versus after the meeting from the same experts. Cues were judged by each expert on each occasion. As the cues differ, residuals of one model would be correlated with predictions of the other, so the Stewart (1978) formula was used (with three C terms). While their post-meeting recommendations are most correlated with their changed post-meeting cue judgments about the features of the proposals, they are still somewhat correlated with their former pre-meeting views.

A Brunswikian Theory of Expert Judgment and its Application to Probabilistic Forecasting

Fergus Bolger
Anglia Ruskin University

Contact: fbolger42@gmail.com

Expert input to forecasting is often needed in data-poor or non-stationary environments. Accurate measurement of the considerable variation in expert substantive and normative abilities is crucial for successful selection and weighting: a theory of expertise should help us develop such measures. I propose that expert forecasters have learned environmental cue-criterion relationships (expressed as linear models). Substantive expertise is operationalised as degree of correspondence between environment and judgment models, while normative expertise is the ability to represent and report this correspondence. Both kinds of expertise are influenced by interactions between environmental (e.g., availability of predictive cues) and personal factors (e.g., good memory). I report the results of simulations where aspects of the environment and experts are manipulated, and probabilistic forecasting performance on realistic datasets is measured.

Choose your Words Wisely: Semiotic Characteristics Predict Halo Effects and Attrition in the Serial Reproduction of Person Descriptions

Karolin Salmen & Klaus Fiedler
Heidelberg University

Contact: karolinsalmen@gmail.com

Descriptions of others frequently reach us through communication, often with several intermediaries. This changes the original information: some pieces are lost (attrition), others are added (halo effect). Prior research attributed these changes to cognitive factors, while this empirical study introduces a novel perspective: A semiotic approach to serial reproduction, which combines the lens model perspective (Brunswik) with a consideration of the verbal sign system used in communication. Two pre-tests (N = 182) established sets of personality traits - grouped by valence and agency/communion - and corresponding behaviours (cues) that vary the semiotic characteristics similarity and cue overlap. In the main study (N = 117), participants used each of the sets to pass on person descriptions in four generations of reproduction. Data were analysed with mixed-effects modelling and signal-detection theory.

Judgements and Decision-Making on Prisoners' Dangerousness: An SJT Perspective

Marcel Guéridon & Susann Prätor
Criminological Service Unit of Lower-Saxony, Germany

Contact: Marcel.Gueridon@justiz.niedersachsen.de

Judgements about "dangerousness" lie at the heart of many prison-related decisions (treatment assignment, security measures, conditional release …). These judgements represent a fundamental, challenging and highly "political" predictive task regarding an inherently probabilistic concept – whose structure may lead to defensive decision-making. Moreover, these judgements can be related to later actual reoffending. Unfortunately, criminological and psychological research often focuses on predictive accuracy of clinical and statistical judgements. This unnecessarily narrow perspective can be extended meaningfully by Brunswikian principles, Social Judgement Theory and extensive analyses of task structure. We will present two applications on judgements and decision-making by prison staff and one related study on judgements by laypersons. In contrast to the usual, predictive focus in criminological research, merits of descriptive research using *SJT* will be highlighted.

Ecological Validity and the "Real World" in the Field of Social Attention

Gijs A. Holleman
Utrecht University

Contact: g.a.holleman@uu.nl

In the field of social attention, there has been a recent surge (2006-present) in papers advocating to design experiments with 'more' or 'higher' ecological validity (Shamay-Tsoory & Mendelsohn, 2019). This argument is certainly not new, nor has this argument been limited to the field of social attention for that matter, as pointed out in detail by Hammond (1998). I will present a brief summary of two articles written by me and my colleagues (Holleman et al., 2020a, 2020b) in which we aimed to inform researchers about the history behind the term 'ecological validity' (Brunswik, 1949; Hammond & Stewart, 2001), and discuss some of the conceptual problems associated with the indiscriminate use of this concept by researchers nowadays to evaluate whether study results generalize to the 'real world'.

Representative Design in Psychological Assessment: A Case Study Using the Balloon Analogue Risk Task (BART)

Markus D. Steiner & Renato Frey
University of Basel

Contact: markus.steiner@unibas.ch

We investigate the role of representative design in achieving reliable and valid psychological assessments by focusing on the Balloon Analogue Risk Task (BART). We demonstrate that the task's original implementation violates the principle of representative design, and show in two studies (N = 772 and N = 632) that participants acquired more accurate beliefs in an adapted, more representative BART. Yet, improving representativeness was insufficient to enhance the task's psychometric properties. We therefore argue that valid task designs may require novel ecological assessments, to identify those real-life behaviors and associated psychological processes that laboratory tasks are supposed to capture and generalize to.

The Ecology of Competition: A Theory of Risk-Reward Environments in Adaptive Decision Making

Timothy J. Pleskac
University of Kansas

Larissa Conradt, Christina Leuker, & Ralph Hertwig
Max Planck Institute for Human Development

Contact: pleskac@ku.edu

Decisions that are made under uncertainty are a vexing problem as one typically seeks to choose between prospects based on the desirability of the possible outcomes and the likelihood of their occurrence. However, in these situations the likelihoods are unknown. One adaptive solution for this problem is to use the risk-reward heuristic inferring that probabilities are inversely related to the magnitude of the payoffs. However, a heuristic is only adaptive if it relies on a frequent and recurrent ecological relationship. While risk-reward relationships are an established property of economic markets it is unclear to what extent they are present in other environments. Here we present the competitive risk-reward ecology theory (CET) that establishes how competition for limited resources is a sufficient ecological mechanism to couple reward size with the probability of success. The result is a consequence of a well-known ecological model known as the ideal free distribution of competitors. According to this model, the number of competitors in a resource patch is proportional to the total amount of resources. We show that this property implies a risk-reward relationship and establishes important boundary conditions for the relationship. For instance, CET predicts degradation in the risk-reward relationship when the system is out of equilibrium. Moreover, it predicts how a range of factors distorts the relationship. For instance, computational limitations among the competitors lead to a shallower risk-reward relationship and smaller rewards being tied with a larger range of probabilities. We will show how human data on their beliefs in the risk-reward relationship reflects these properties. In sum, grounding people's inferences in CET demonstrates how the behaviors of a boundedly rational mind can be better predicted once accounts of the mind and the environment are fused.

Precise/Not Precise (PNP): A Brunswikian Model that Uses Judgment Error Distributions to Identify Cognitive Processes

Joakim Sundh, August Collsiöö, Philip Millroth, & Peter Juslin
Uppsala University

Contact: august.collsioo@psyk.uu.se

In 1956, Brunswik proposed a definition of what he called intuitive and analytic cognitive processes, not in terms of verbally specified properties, but operationally based on the observable error distributions. In the decades since, the diagnostic value of error distributions has generally been overlooked, arguably because of a long tradition to consider the error as exogenous (and irrelevant) to the process. Based on Brunswik's ideas, we develop the precise/not precise (PNP) model, using a mixture distribution to model the proportion of error-perturbed versus error-free executions of an algorithm, to determine if Brunswik's claims can be replicated and extended. In Experiment 1, we demonstrate that the PNP model recovers Brunswik's distinction between perceptual and conceptual tasks. In Experiment 2, we show that also in symbolic tasks that involve no perceptual noise, the PNP model identifies both types of processes based on the error distributions. In Experiment 3, we apply the PNP model to confirm the often-assumed "quasi-rational" nature of the rule-based processes involved in multiple-cue judgment. The results demonstrate that the PNP model reliably identifies the two cognitive processes proposed by Brunswik, and often recovers the parameters of the process more effectively than a standard regression model with homogeneous Gaussian error, suggesting that the standard Gaussian assumption incorrectly specifies the error distribution in many tasks. We discuss the untapped potentials of using error distributions to identify cognitive processes and how the PNP model relates to, and can enlighten, debates on intuition and analysis in dual-systems theories.

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

http://www.brunswik.org/