THE BRUNSWIK SOCIETY NEWSLETTER

VOLUME 37

NOVEMBER 2022

ISSN 2296-9926

ABOUT

The Brunswik Society is an international association of researchers who are interested in understanding and improving human judgment and decision making. Founded in 1985 by Kenneth Hammond, members of the Society share an appreciation for the work of psychologist Egon Brunswik.

STAY CONNECTED

Visit the Brunswik Society website at https://brunswiksociety.org/.

To receive information about the Society's annual meeting and newsletter, email info@brunswiksociety.org.

Follow us on Twitter! @BrunswikThe

Imprint

The Brunswik Society Newsletter, ISSN 2296-9926 is published yearly

Editorial staff: Adaryukov, J., Collsiöö, A., Hamm, R. M., Kaufmann, E., & Molinaro, K.A.

TABLE OF CONTENTS

FOREWORD

We are very pleased to present the 2022 Brunswik Society Newsletter, many thanks to all authors for their contributions!

In 2015, we commemorated the passing of Kenneth Hammond, who reinvigorated Brunswikian research. Now, seven years later, we mourn another friend. Robert "Bob" Wigton was an inspiration to many, a multitalented researcher who "epitomized the concept of the Renaissance man" (Tape). One of our society's longest-term members, he contributed constantly to the newsletter and the meeting for two decades. His absence in these pages and in our company is keenly felt. However, Bob's memory endures through those he inspired and will continue to inspire long after his passing. Therefore, this newsletter is not a funeral, but a celebration of life, both that of Bob and that of the ideas he championed.

Indeed, this issue represents just how far Brunswikian decision-making research has come since this society's inception. On the methodological level, our contributors propose multiple new ways of thinking about the lens model and representative design. From new methods of parsing different sources of judgment error (Bago d'Uva and O'Donnell), to a tool for verifying ecological validity in research designs (Naumann et al.), to enhancing the lens model framework through machine learning (Binz) and application to actions rather than to judgments (Hamm), readers will see the continued evolution of how we approach judgment and decision-making.

To what ends? Our contributors have answers for this as well. Some follow the same path as Bob and apply Brunswikian theory to the medical field (Papa; Waghorn, Rakow and Stevenson). Others examine how this research can apply to new technological questions, such as decision-making with humans and AI (Luna-Reyes and Harrison) and predicting the risk of a cyber-attack (Pitterele). Tarlao and Gustavson analyze the representativeness of soundscape evaluation, comparing how people evaluate soundscapes in the laboratory versus in their natural settings. Meanwhile, Kaufmann draws attention to how the lens model could apply quite well to education research. These are only some of the contributions highlighting the ever-expanding possibilities for what Brunswikian analysis can do.

All of these possibilities would amount to little, however, without people pursuing them. Our contributions from doctoral students, including those by Bolesta and Conlon, affirm that there will be people to carry them forward. Furthermore, we are continuing our work of making this research more accessible to all. We outline in this newsletter the BrunsWiki project, calling interested writers to contribute to Wikipedia entries for Brunswikian concepts and ensure key ideas are not lost. Through our own continued archiving work and the interest of new researchers, this area of research will not only stay alive but continue to flourish.

With all these explorations by different groups of researchers, we note some variation in the terminology attributed to the Brunswikian framework, such as the "lens model" and "ecological validity". First, we acknowledge that the multiple regression "lens model equation" was developed by Hammond and colleagues and by Tucker, during the decade after Brunswik's death, and that multiple lenses may be found in Brunswik's framework (a pinhole camera model connecting

perception through the organism to action; a convex glass lens model connecting objects in the world through perceptible cues to a cognitive representation of those objects) so "lens model" does not equal "application of the lens model equation." Second, there is some variation in the use of the term "ecological validity." In the lens model equation context, ecological validity refers to the correlation between a particular cue and the object in the ecology, while in the psychology field broadly it is applied to the relation between a study's conditions and the ecology, such that the results of a study may be generalized. In Brunswikian literature that is called "representativeness." The history of Gibson's and Brunswik's use of ecological validity was explored by John Kihlstrom in the 2020 Brunswik Society Newsletter. The editors discussed these various terminology issues with the contributors, in the spirit of welcoming all who venture into exploring these ideas.

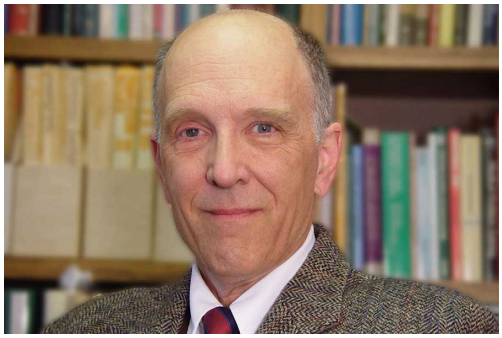
One person could not have done all this alone. This issue, and this newsletter more broadly, represents the work of hundreds, each contributing to an even more robust understanding of our actions and our judgments. Whether you have contributed to the newsletter directly, attended our annual meetings, or even read through these contributions, you are part of that work as well, and we thank you for it. To Bob and other colleagues departed, we thank you as well. Past, present, future, or all of the above, this newsletter is dedicated to you.

Sincerely,

James Adaryukov, August Collsiöö, Robert M. Hamm, Esther Kaufmann, and Kylie A. Molinaro

IN MEMORIAM

Robert Swift Wigton, MD, MS



January 22, 1942 – August 22, 2022

With the recent passing of Robert "Bob" Wigton, the Brunswik Society lost one of its long-time active members. Although Bob had been dealing with a chronic form of leukemia for many years, he only recently retired from his faculty position at the University of Nebraska Medical Center (UNMC) where he devoted his entire research career to study medical education and to study judgment and decision making in various medical settings. In his later years, he served as the "unofficial" UNMC historian and became the principal donor to the Wigton Heritage Center (https://www.unmc.edu/wigton/index.html) on the UNMC Campus.

Bob grew up in Omaha, Nebraska where his father practiced psychiatry and neurology. Multiple ancestors on both sides of his family practiced medicine as documented in this excellent online exhibit: https://wigton.unmc.edu/LegaciesHealing. Bob graduated from Harvard University with a B.A. in English Literature and then returned to Nebraska for his medical studies where he received his M.D. in 1969. He stayed on as an internal medicine resident and chief resident while also pursuing a master's degree in physiology. He joined the medical faculty at Nebraska in 1974 with an initial research focus in medical education. Bob was an early pioneer in promoting the use of micro-computers (early desktop computers) in medical education. During the 1980s and 1990s, he coordinated the microcomputer users' interest groups for both the Society for Medical Decision Making and the Society of General Internal Medicine. He also sponsored an annual national conference at UNMC called *Micro-computers in Medical Education*.

In 1982, Bob did a sabbatical at the University of Pennsylvania where he developed research expertise in the psychology of medical judgment and decision making. His earliest judgment work used conjoint analysis (see Green & Rao, 1971), but he would later focus on policy capturing using fractional factorial designs of Plackett and Burman (1946). I met Bob in 1985 at the annual meeting of the American Federation of Clinical Research (AFCR) in Washington, DC and was excited to be recruited by him to UNMC, where we subsequently enjoyed a long collaboration in judgment research.

Ken Hammond & Bob Wigton, 1999

During the mid-1980s, Bob became aware of Professor Ken Hammond's contributions to Brunswikian research. Since Bob's vacation cabin in Colorado was not far from Ken Hammond's home in Boulder, Bob began to periodically visit Ken when Bob was in Colorado. Bob took me along for several of those trips and I recall pleasant afternoons in the Hammonds' back yard discussing drafts of Ken's 1996 book. Ken encouraged both of us to start attending the annual Brunswik meetings which provided an additional professional home for Bob besides the Society for Medical Decision Making and the Society for General Internal Medicine. Bob became a regular presenter at the Brunswik meetings and a regular contributor to the Brunswik Newsletter. Esther Kaufmann searched the archives and found that Bob made contributions to the Brunswik Newsletter virtually every year from 1991 to 2015. Bob published more than 160 articles, most concerning medical judgment and decision making. His classic article, *Use of Linear Models to Analyze Physicians' Decisions* (Wigton, 1988) provides an excellent overview of the various approaches that he and others applied to understanding medical judgment. He and his collaborators developed clinical prediction rules for a number of medical conditions, including pharyngitis, pulmonary embolism, urinary tract infection, and pneumonia. He led multiple short courses on linear judgment models at the meetings of the Society for Medical Decision Making starting in the 1980s and intermittently (up through) the 2014 European Medical Decision Making meeting in Antwerp, Belgium.

Bob with his painting of Chimney Rock, Nebraska, 2016

Bob Wigton's expertise and interests encompassed numerous fields besides judgment psychology. Over the years, these included art history, making art, performing jazz, dirt motorcycles, botany, gardening, science fiction, medical history, digital photography, and exploring nature.

While in college, he developed skills in cartooning which he later used in his teaching to enrich his lectures. When announcing an upcoming talk, he would often include a caricature of the speaker. After his retirement from clinical medicine, he further developed his artistic abilities in oil paint and watercolors. Nebraska landscapes became a favorite subject for his art, as did paintings of various places he visited on his travels. He showed his art in an Omaha gallery.

Bob had a particular love for nature. Throughout the years, he and his wife, Debbie, loved to hike and bird watch in the forests along the Missouri River near Omaha. They often vacationed in Rocky Mountain National Park where Bob and his family had been visiting since he was young. His routine was to hike every day, weather permitting, progressing to higher and more challenging routes as he became acclimated to the altitude. Bob also enjoyed fly fishing, photography, and bird watching.

Fly fishing in Rocky Mountain National Park, 1999

Since about the year 2000, Bob became passionate about digital photography. He was fascinated by the possibilities of the digital darkroom, and he became expert in using Adobe Photoshop to achieve stunning results. He always took a camera (or two) on his hikes and especially loved to zoom in on wildflowers and wildlife. Examples of his photography can be seen throughout the corridors of the UNMC campus. His camera was also ubiquitous when he traveled to scientific meetings. He enjoyed giving talks about the places he visited, using his photographs to illustrate the history and culture of his travel destinations.

Photographing a turtle, 2002

During the later phase of his long and productive career, he focused his energies on studying the history of medicine in Nebraska and preserving the history of his own institution, UNMC. He worked with archival librarians to digitize and catalog artifacts as well as to obtain oral histories from long-time members of the Omaha medical community. These efforts culminated in building the Wigton Heritage Center, which was dedicated in 2021. Its exhibits, both physical and on-line, serve as a lasting tribute to Bob's passion for history and preservation.

Bob Wigton epitomized the concept of a "Renaissance Man." This brief memory of his life does not do justice to his many talents. A quotation from a 1914 eulogy for his great-grandfather, Alonzo Wigton, is apt here as we remember Bob: "He measured success by what he gave to the world rather than by what he took from it."

— Tom Tape, MD; Emeritus Professor, University of Nebraska Medical Center

References

Green, P. E., & Rao, V. R. (1971). Conjoint Measurement-for quantifying judgmental data. *Journal of Marketing Research*, *8*(3), 355–363.

Plackett, R. L., & Burman, J. P. (1946). The design of optimum multifactorial experiments. *Biometrika*, *33*(4), 305–325.

Wigton, R. S. (1988). Use of linear models to analyze physicians' decisions. *Medical Decision Making*, *8*(4), 241–252.

Key Work Related to Brunswik

Engel, J. D., Wigton, R., LaDuca, A., & Blacklow, R. S. (1990). A social judgment theory perspective on clinical problem solving. *Evaluation & the Health Professions*, *13*(1), 63–78.

Tape, T. G., Heckerling, P. S., Ornato, J. P., & Wigton, R. S. (1991). Use of clinical judgment analysis to explain regional variations in physicians' accuracies in diagnosing pneumonia. *Medical Decision Making*, *11*(3), 189–197.

Wigton, R. S. (2008). What do the theories of Egon Brunswik have to say to medical education? *Advances in Health Science Education*, *13*, 109–121.

Wigton, R. S. (1996). Social judgement theory and medical judgement. *Thinking & Reasoning*, *2*(2–3), 175–190.

Additional work can be found through his Google Scholar profile here: https://scholar.google.com/citations?hl=en&user=7UtNfsAAAAAJ&view_op=list_works%20%5bscholar.google.com%5d

SOCIETY NEWS

28th International (Virtual) Meeting of the Brunswik Society

Free event – register your name, affiliation, and email address with esther.kaufmann@gmx.ch

8th December 2022, 12.00-14.15 EST (17.00-19.15 GMT) via Zoom

Opening Remarks – Mandeep Dhami (Middlesex University, London, UK)

Title: News from the Brunswik Society
Presenters: Karolin Salmen, Gijs Holleman, Mandeep Dhami, & Esther Kaufmann

Title: Generative rationality, growing awareness and uncertain ecologies
Presenter: Teppo Felin (Utah State University, US)
Abstract: Bounded and ecological rationality build on a view of perception that is anchored on veridicality, statistics, and an all-seeing view of reality. I address the underlying assumptions of this work, from psychophysics and computer vision to more modern work on humans as intuitive statisticians. I argue for a generative view of rationality—highlighting how awareness is "grown" by organisms and humans top-down. I offer examples from evolutionary biology and modern decision environments, and contrast these with common examples in the bounded rationality literature. I conclude with a plea for more work on the organism-, task- and problem-specific contingencies of both perception and rationality.

Title: How experimental methods shaped views on human competence and rationality
Presenter: Tomas Lejarraga (Max Planck Institute for Human Development, GER)
Abstract: In the early 1970's, behavioral decision research underwent a dramatic change. In 1967, an exhaustive review by Peterson and Beach provided experimental evidence for Brunswik's ideas. The review showed that people could be viewed as intuitive statisticians. But in 1974, Tversky and Kahneman rejected this conclusion, arguing that people rely on a limited number of heuristics that work well most of the time but are prone to bias. Their heuristics-and-biases research program changed how scientists view the mind and established an experimental protocol that relied on described scenarios rather than learning and experience. We examine lines of research in the intuitive-statistician research program and that spurred by the heuristics-and-biases program and examine how the focus on description at the expense of learning has shaped the influential view of the error-proneness of human cognition.

Title: Using Brunswik's lens model to explain judgment inaccuracy through extended decomposition of the Brier Score
Presenters: Teresa Bago d'Uva (Erasmus School of Economics and Tinbergen Institute, NL) & Owen O'Donnell (Erasmus School of Health Policy and Management, NL)
Abstract: Identification of sources of judgement inaccuracy is critical to improving judgements and decisions. We use Brunswik's lens model to make an extended decomposition of the Brier Score measure of inaccuracy – the Mean Squared Error – into a rich array of potential sources of error in judgements. This goes beyond the focus of some other decompositions on correlation, separates out the contribution of noise and inappropriate weighting of information contained in measured cues, and reveals the offsetting effect of private information not captured by those cues. We demonstrate the usefulness of this approach in explaining accuracy of professional economic forecasters and laypersons predicting their own longevity. Application to many other expert and layperson judgements can potentially help identify sources of error that lead to suboptimal decisions.

Discussion Paper in Memory of Robert Wigton: Robert Wigton's contributions to Brunswikian research and teaching in medicine

Presenter 1: Thomas Tape (University of Nebraska Medical Center, US)
Abstract 1: Doctor Robert (Bob) Wigton's research focused on medical education, especially understanding the factors that physicians and trainees considered when making diagnostic assessments. He explored the concepts espoused by Egon Brunswik and championed by Ken Hammond and applied them to the study of medical judgment across a variety of diagnostic problems. This survey of Bob Wigton's most important papers will describe the evolution of his research approaches from conjoint analysis to Brunswikian linear models as well as his study of feedback approaches to improving medical judgment.

Presenter 2: Robert M. Hamm (University of Oklahoma Health Sciences Center, US)
Abstract 2: Bob Wigton's teaching of Brunswikian research focused on getting physicians to know enough to actually do lens model studies. To make it practical, he simplified the methods from the lens model equation ideal. He did not bother addressing the $C*sqrt(1-R_e^2)*sqrt(1-R_j^2)$ portion of the lens model equation, and he used fractional factorial designs rather than having stimulus sets in which the cues correlated with each other as is found in the ecology. He steadily presented these ideas over several decades and collaborated with others to use the methodology to reveal interesting, sometimes depressing, facts about physicians' clinical judgment accuracy.

Closing Remarks – Tom Stewart (University at Albany, US)

Social/Networking Hour!

8th December 2022, starts 14.30 EST (19.30 GMT)

Free event – invites will be sent to meeting delegates

The BrunsWiki Challenge: Your Society Needs You!

Mandeep K. Dhami

Middlesex University London, United Kingdom

✉ m.dhami@mdx.ac.uk

Egon Brunswik's and Kenneth R. Hammond's contributions were vast in terms of coverage of theoretical, methodological and applied issues. Together, their contributions also spanned decades and scientific fields. This has posed challenges for researchers who wish to understand their ideas and impact, especially in a holistic way.

Over the decades, Brunswik's and Hammond's many contributions have been usefully collated into edited volumes (e.g., Brehmer & Joyce, 1988; Cooksey, 1996; Hammond, 1966; Hammond & Stewart, 2001; Hammond, & Wascoe, 1980; Rappoport & Summers, 1973; special issue of *Thinking and Reasoning*, 1996). Some ideas have also been the subject of in-depth reviews (e.g., Dhami, Hertwig & Hoffrage, 2004; Dhami & Mumpower, 2018; Dhami & Olsson, 2008; Karelaia, & Hogarth, 2008; Kaufmann & Athanasou, 2009; Kaufmann, Reips, & Wittmann, 2013). However, as time as gone by, many of these books have become difficult to access, and whereas the review articles are more easily accessible, they are necessarily focused in their coverage, thus leaving some important ideas "buried".

Valiant efforts have also been made by Thomas R. Stewart to make Brunswik's original work available through the Brunswik Society webpage. Tom and Karolin Salmen have recently refreshed the website and moved it to a new link (see https://brunswiksociety.org/).

While this is a great, easy-to-use repository, it is does not (yet) feature all of Hammond's contributions. In addition, websites are not always easily located by those who many need to know the information they contain (e.g., a Google search for "lens model" will not necessarily bring the Society's website to the fore).

Therefore, in January 2022, I set the Brunswik Society a challenge – namely of establishing Wikipedia entries for key theoretical and methodological concepts associated with the work of Brunswik and Hammond, as well as entries for key figures such as Berndt Brehmer whose contributions significantly extended work in the Brunswik-Hammond tradition. Table 1 presents a list of the "BrunsWiki" entries we want to develop and publish.

The final column in Table 1 shows the members of the Society who have kindly volunteered to help write first drafts of entries – several of whom have come out of retirement simply to do so. The Society and I are very grateful to them. If you are interested in helping to either write an entry or review a draft entry, then please contact me at m.dhami@mdx.ac.uk. Also, we are in desperate need of help from someone who has experience with publishing on Wikipedia – if this is you, then please contact me and I will put you in touch with Gijs Holleman, who is graciously spending time learning how to do this on our behalf while also moving to his first faculty position. We hope to have all entries finalized and uploaded onto Wikipedia before the 2023 annual meeting.

Table 1. Proposed BrunsWiki entries.

Topic	Current Status	First Draft Volunteer
Leading Figures (deceased)		
Egon Brunswik	Needs updated: https://en.wikipedia.org/wiki/Egon_Brunswik There is an existing German article: https://de.wikipedia.org/wiki/Egon_Brunswik	STEWART
Kenneth R. Hammond	Does not exist	STEWART
Berndt Brehmer	English article: Does not exist There is an existing, rudimentary Swedish article: https://sv.wikipedia.org/wiki/Berndt_Brehmer	???
Main Ideas		
Probabilistic functionalism	Does not exist	???
Representative design	Does not exist	DHAMI
Lens model	Does not exist	KAUFMANN
Judgment analysis	There is an existing article on policy capturing that needs some adaptation. We can edit that and point to judgment analysis. https://en.wikipedia.org/wiki/Policy_capturing	STEWART
Lens model equation	Does not exist	HAMM
Ecological validity	Needs updated with a stronger link to Brunswik's definition: https://en.wikipedia.org/wiki/Ecological_validity There is also an article on ecological validity (perception): https://en.wikipedia.org/wiki/Ecological_validity_(perception)	???
Vicarious functioning	Final version needs published	ADELMAN
Social judgment theory	Needs updated; is conflated with theory of social persuasion: https://en.wikipedia.org/wiki/Social_judgment_theory	???
Multiple cue probability learning	Does not exist	HOLZWORTH
Interpersonal learning	Does not exist	MUMPOWER
Interpersonal conflict	Does not exist	DHAMI
Cognitive feedback	Does not exist	???
Cognitive continuum theory	Does not exist	DHAMI

Idiographic v. nomothetic	Needs updated: https://en.wikipedia.org/wiki/Nomothetic_and_idiographic	???

References

Brehmer, B., & Joyce, C. R. B. (1988). (Eds.). *Human judgment: The SJT view*. New York: North-Holland.

Cooksey, R. W. (1996). *Judgment analysis: Theories, methods, and applications*. New York: Academic Press.

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin, 130*(6), 959–988.

Dhami, M. K., & Mumpower, J. (2018). Kenneth R. Hammond's contributions to the study of human judgment and decision making. *Judgment and Decision Making*, *13,* 1–22.

Dhami, M. K., & Olsson, H. (2008). Evolution of the interpersonal conflict paradigm. *Judgment and Decision Making*, *3*(7), 547–569.

Hammond, K. R. (1966) (Ed.). *The psychology of Egon Brunswik*. New York: Holt, Rinehart, & Winston.

Hammond, K. R., & Stewart, T. R. (2001). (Eds.). *The essential Brunswik*. Oxford: Oxford University Press.

Hammond, K. R., & Wascoe, N. (1980). (Eds.). *Realizations of Brunswik's representative design.* San Francisco: Jossey-Bass.

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: a meta-analysis of lens studies. *Psychological Bulletin*, *134*(3), 404–426.

Kaufmann, E., & Athanasou, J. A. (2009). A meta-analysis of judgment achievement as defined by the lens model equation. *Swiss Journal of Psychology / Schweizerische Zeitschrift für Psychologie / Revue Suisse de Psychologie*, *68*(2), 99–112.

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), e83528. https://doi.org/10.1371/journal.pone.0083528

Rappoport, L., & Summers, D. (1973) (Eds.), *Human judgment and social interaction*. New York: Holt, Rinehart, & Winston.

Special issue of *Thinking & Reasoning* (1996), vol 2, issue 2–3.

CONTRIBUTIONS

Using Brunswik's Lens Model to Explain Judgement Inaccuracy Through Extended Decomposition of the Brier Score

Teresa Bago d'Uva

Erasmus School of Economics and Tinbergen Institute, Netherlands

✉ bagoduva@ese.eur.nl

Owen O'Donnell

Erasmus School of Economics and Tinbergen Institute & Erasmus School of Health Policy and Management, Netherlands

Identification of sources of judgment inaccuracy is critical to improving judgments and decisions. In a recent paper (Bago d'Uva & O'Donnell, 2022), we use Brunswik's (1952) lens model to make an extended decomposition of the Brier (1950) Score measure of inaccuracy – the mean squared error (MSE) – into a rich array of potential sources of error in judgments.

Our approach offers three advantages over the lens model equation (Tucker, 1964). First, like other lens model decompositions of judgment MSE (Kane & Broomell, 2020; Lee & Yates, 1992), it goes beyond examination of correlation between outcomes and judgments to explain the *magnitude of judgment errors*. Second, it identifies the contribution to inaccuracy of variation in judgments that is not explained by the outcome, which, following Yates (1982), we label *noise*. This can help identify differences in judgment accuracy that arise from differences in ability to report judgments that respond consistently to cues and ignore irrelevancies. Third, the extended decomposition does not only identify the aggregate contribution to inaccuracy of divergence between weights given to cues in models of the outcome and the judgments, but it also identifies separate contributions of inappropriate weights on specific cues.

Furthermore, our approach goes beyond prior lens model decompositions of judgment accuracy (Kane & Broomell, 2020; Stewart & Lusk, 1994) by identifying the contribution of relevant, supplementary, and utilized information. While the full lens model equation allows for correlation between the residuals of the outcome and judgment models (often denoted C), this has been presumed to mostly reflect unmodelled nonlinear cue-outcome and cue-judgment relationships that can be reduced by refining the specification. Namely, it has been omitted from the MSE decompositions. However, when the outcome is personal – a health event, for example – the past and likely future behavior of the judge gives them relevant information that is not measured in the data and so cannot be used to improve model specification. The quantity and distribution of this *private information* is particularly important to the operation of markets (Akerlof, 1970; Rothschild & Stiglitz, 1976).

Our extended decomposition of the Brier Score identifies the contributions to judgment inaccuracy of 1) judgment task difficulty, 2) outcome predictability, 3) bias, 4) inappropriate weighting of cues, 5) private information, and 6) noise. The figure shows how this decomposition compares with the covariance and Yates MSE decompositions, and the use it makes of the lens model.

We use the decomposition to explain the accuracy of professional forecasters predicting economic recession (N = 76 forecasters, 3810 judgment-outcome pairs, mean Brier Score = 0.099) and laypersons predicting their own longevity (N = 4946 individuals, Brier Score = 0.246). In both applications, judgment difficulty makes the largest contributions to inaccuracy and its variation, although this is partially offset by outcome predictability. Inappropriate weighting of cues is substantial and also helps explain variation in inaccuracy. In the longevity application, low discriminatory power of the judgments is partly due to insufficient responsiveness to mortality risk factors, particularly among the least educated. In both applications, noise in judgment residuals (net of predictions from cues) is an important source of inaccuracy and its variation. This noise is partially offset by non-negligible (private) information these residuals contain on the outcome.

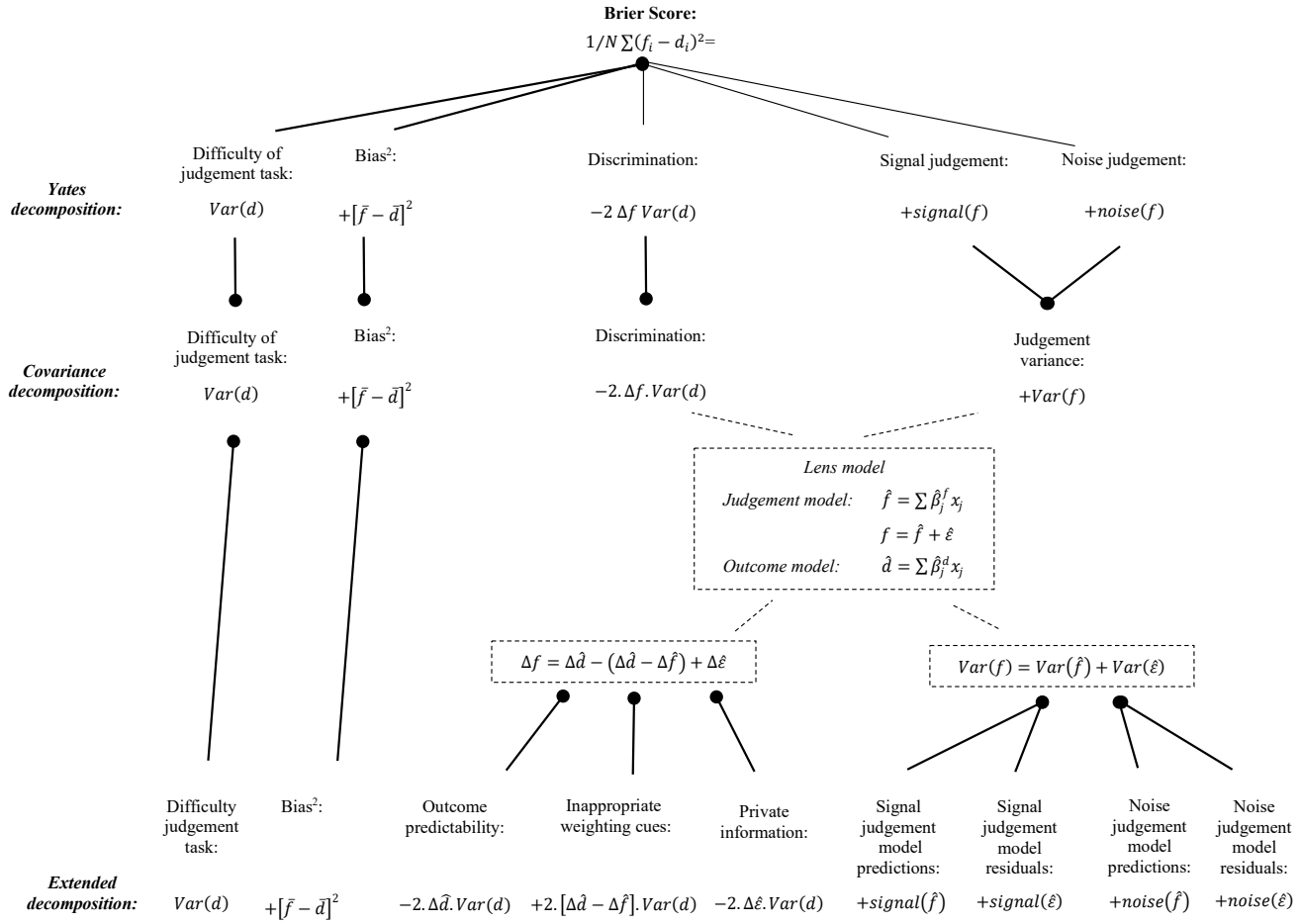


Figure 1. Covariance, Yates, and Lens Model Extended Decompositions of Brier Score from Bago d'Uva and O'Donnell (2022).

References

Akerlof, G. (1970). The market for 'lemons': Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*(3), 488–500.

Bago d'Uva, T., & O'Donnell, O. (2022). Explaining probability judgment inaccuracy: A lens model extended decomposition of the Brier Score. *Decision*, *9*(1), 74–90.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, *78*, 1–3.

Brunswik, E. (1952). *The conceptual framework of psychology.* Chicago: University of Chicago Press.

Kane, P. B., & Broomell, S. B. (2020). Applications of the bias–variance decomposition to human forecasting. *Journal of Mathematical Psychology*, *98*, 102417.

Lee, J., & Yates, J. F. (1992). How quantity judgment changes as the number of cues increases: An analytical framework and review. *Psychological Bulletin*, *112*(2), 363–377.

Rothschild, M., & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, *90*(4), 629–49.

Stewart, T. R., & Lusk, C. M. (1994). Seven components of judgmental forecasting skill: Implications for research and the improvement of forecasts. *Journal of Forecasting*, *13*(7), 575–599. https://doi.org/10.1002/for.3980130703

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.

Yates, J. F. (1982). External correspondence: Decompositions of the mean probability score. *Organizational Behavior and Human Performance*, *30*(1), 132–156.

Heuristics from Bounded Meta-Learned Inference

Marcel Binz

Max Planck Institute for Biological Cybernetics, Germany

✉ marcel.binz@tuebingen.mpg.de

Imagine having to decide which of two movies you are going to watch tonight: Movie A versus Movie B. Movie A has a higher average rating on a website that you trust, while Movie B is directed by a known director and has previously won an Oscar for the best picture. From past experiences, you know that rating is the best indicator of a good movie. Whether the movie won an Oscar and who directed it is less important for how much you normally enjoy watching a movie. How do people make decisions like this?

The question of how people decide between two options is as fundamental as its answer is contentious. Indeed, even though we make countless such decisions every day, the underlying principles of these decisions are still debated in psychology (Todd & Gigerenzer, 1999), behavioral economics (Samuels et al., 2012), and neuroscience (Camerer et al., 2005). Traditionally, researchers have approached this problem by looking at how rational agents decide. From this ideal observer perspective, it is assumed that people weigh different attributes of each option appropriately to combine information from all available sources. Psychologists were however quick to point out that rational decision-making can be too burdensome (Simon, 1990; Tversky & Kahneman, 1974). Instead, they suggested that human decision-making may be based on a variety of heuristics, which are simple strategies that ignore part of the relevant information.

Heuristics are computationally efficient decision-making strategies that can be surprisingly competitive in many real-world benchmarks (Czerlinski et al., 1999). These properties have often been used to justify them as models of human decision-making. Yet, strong evidence for heuristics in human decision-making is still lacking as empirical studies have often produced mixed results.

In our recent paper (Binz et al., 2022), we proposed a novel computational model that advances our understanding of heuristic decision-making by explaining how different heuristics are discovered and how they are selected. This model – called bounded meta-learned inference (BMI) – is based on the idea that people make environment-specific inferences about which strategies to use while being efficient in terms of how they use computational resources. BMI discovers decision-making strategies through a resource-rational algorithm (Gershman et al., 2015; Lieder & Griffiths, 2019) that has been adapted to an environment over time via meta-learning (Schmidhuber et al., 1996; Thrun & Pratt, 1998). More specifically, it specifies a family of learning algorithms that attempt to achieve optimal behavior subject to the constraint that they can be implemented with a given number of bits. Like ideal observer models, BMI attempts to infer optimal decision-making strategies but does so while taking computational resources into account. Like heuristics, strategies inferred through BMI are tailored to a specific environment. However, unlike heuristics, the inductive biases of such strategies have been discovered based on previous interactions with the environment instead of being predetermined by researchers.

Through a series of model simulations, we show that BMI discovers several previously suggested heuristics. In particular, we found that:

- If the model knows the correct ranking of attributes but not their weights, then it learns a strategy that makes decisions based only on the attribute with the highest ranking, a form of one-reason decision-making.
- If the model knows that the direction of correlation between attributes and outcome is positive, then it learns a strategy that makes decisions based on equal weighting.
- If the model does not know either the ranking or the direction of attributes, it learns to use individual weights for each attribute.

These model simulation results enabled us to make precise predictions about when to expect heuristics as part of human decision-making and when not. We subjected these predictions to a rigorous test in three paired comparison experiments and found that the vast majority of participants applied decision-making strategies as predicted by BMI (see Figure 1 for a summary).

Taken together, our work demonstrates that people apply heuristics whenever they are optimal strategies for the current environment after considering limited computational resources. It also highlights the interaction between human cognition and the environment it takes place in – one of the concepts that Brunswik strongly argued for. Brunswik's influence is not only limited to the conceptual level but permeates deep into the computational models we have used in our work: BMI infers a linear weighting of features, essentially making it a direct descendant of Brunswik's lens model.

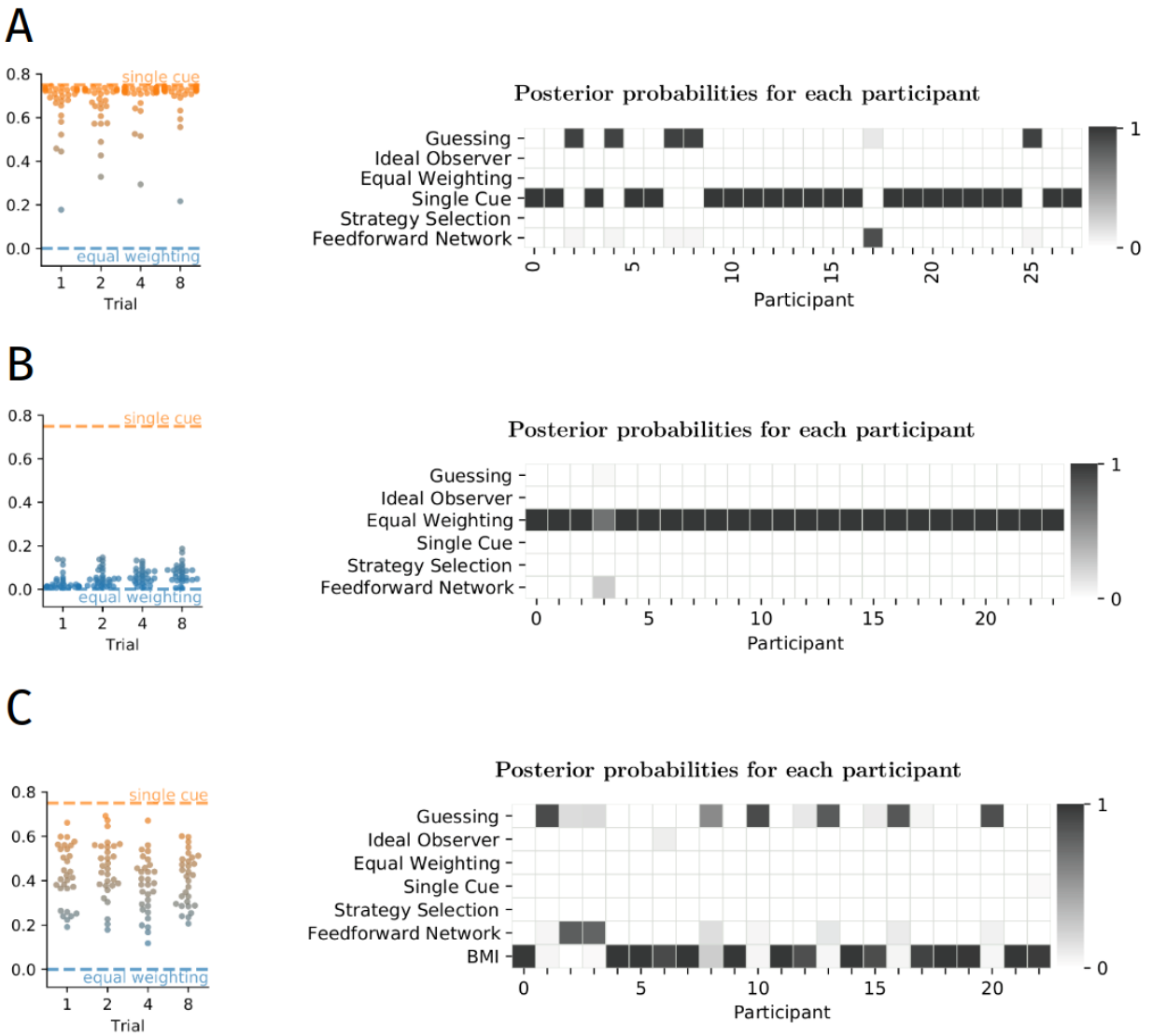


Figure 1. Results from three empirical studies. High values in the left column indicate similarity to a single cue heuristic. Low values indicate similarity to an equal weighting heuristic. **A:** BMI discovered a single cue heuristic as resource-rational solution when given information about the ranking of features. In a subsequent experiment, we confirmed that people predominantly applied this single cue heuristic when they had access to the feature ranking. **B:** BMI discovered an equal weighting heuristic as resource-rational solution when given information about the direction of features. In a subsequent experiment, we confirmed that people predominantly applied this equal weighting heuristic when they had access to feature directions. **C:** When given no information about ranking or direction, BMI makes decisions based on a weighted combination of features. In a subsequent experiment, we confirmed that people also apply this strategy whenever they had no side information about ranking or direction. Figure adapted from Binz et al. (2022).

References

Binz, M., Gershman, S. J., Schulz, E., & Endres, D. (2022). Heuristics from bounded meta-learned inference. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000330

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature*, *43*(1), 9–64. https://doi.org/ 10.1257/0022051053737843

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, & The ABC Research Group. (Eds.), *Simple heuristics that make us smart* (pp. 97–118). Oxford University Press.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press.

Lieder, F., & Griffiths, T. L. (2019). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, E1.

Samuels, R., Stich, S., & Bishop, M. (2012). *Ending the rationality wars*. Collected Papers, Volume 2: Knowledge, Rationality, and Morality, 1978–2010 (pp. 191–223). Oxford University Press.

Schmidhuber, J., Zhao, J., & Wiering, M. (1996). *Simple principles of metalearning*. *Technical Report IDSIA*, *69*, 1–23.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*(1), 1–19.

Thrun, S., & Pratt, L. (1998). Learning to learn: Introduction and overview. In: Thrun, S., Pratt, L. (eds) *Learning to Learn*. Springer

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Contextualizing Political Ideology: On the Impact of Measurement, Domain, and Identity

Deliah Sarah Bolesta

Center for Criminological Research Saxony, Germany

✉ deliah.bolesta@zkfs.de

When Aristotle referred to any human being as a *zoon politikon* he described them as being "by nature a political animal" (Aristotle, trans. 2009, book 1, II). We are naturally sociable, drawn to social communities and as a function thereof, we may naturally develop attitudes towards political matters (Piepenbrink, 2001). Modern theories of political socialization still emphasize on the role of social context in the sense that political attitudes develop within ecological systems (Bronfenbrenner, 1979). They are shaped within micro- (family, peers, educational institutions) and macro-systems (political events, socio-structural features, political systems, political climate), as well as by the media (Eckstein, 2019). Understanding political attitudes as dynamic constructs that develop within the complex interplay of various contextual factors, they may well, once developed, also be affected by those same factors (Jennings, 1990).

Accommodating said context-sensitivity, it was the aim of this dissertation to assess the relationship of one's political ideology to certain psychological underpinnings by leading it "through a context which the world provides" (James, 1975, p. 35). Taking on an ecological perspective along the lines of Brunswik (1955, 1956), a special focus was given to representative sampling. Assuming that an individual's perception and behavior is always organized in reference to and not independent from their environment (Dhami, Hertwig, & Hoffrage, 2004, p. 959), Brunswik suggested that research designs should be representative, in that experimental stimuli may be sampled from within the individual's environment that one wishes to generalize to. That is, as Kessler et al. (2015, p. 31) put it, research should "follow Brunswik's idea of varying the environmental stimuli in order to disentangle psychological processes from content. This would be possible only by varying the content of stimuli either systematically or according to the typical distribution in a certain environment." Taking on an ecological perspective and considering contextual factors as a source of variance, the four lines of research in this dissertation aimed at extending the knowledge on how psychological needs and motives relate to political ideology.

The first project assessed whether cognitive closure could be a function of the political domain in question rather than one's political ideology alone. Political domains used in these studies (religion, climate change, abortion, same-sex marriage, and gun ownership) were sampled with regard to relevant political discussions in the country of participants' origin at the time of data collection. Across all three studies, I consistently found higher need for cognitive closure in liberals relative to conservatives when an environmental domain was addressed. Vice versa, conservatives displayed higher need for cognitive closure than liberals when religion or abortion was addressed.

The second project examined the degree of entanglement between political attitudes (easy vs. hard policy issues; Carmines & Stimson, 1980) and the level of identity (personal vs. political identity; Tajfel & Turner, 1979) dependent on one's political orientation. In two studies, individuals

endorsed counter-attitudinal issues stronger when their personal identity was salient, and pro-attitudinal issues more strongly when their political identity was salient. This was only found for hard issues and for individuals whose "own" party (i.e., the party associated with their own political orientation) was not in government. These results attest to the importance of identity, the nature of issues and societal context in determining people's political attitudes.

A third project investigated the effect of exposure to in- or out-group fake news on political polarization. I presented participants with real-life fake news and orthogonally crossed correction (disclosure vs. no disclosure) and group membership of the sender (in-group vs. out-group vs. ambiguous). Exposure to in- and out-group fake news were associated with high affective polarization (Study 1) while higher levels in attitude polarization were found after exposure to out-group fake news (Study 2). Informing participants about the fake news nature of the post attenuated affective polarization for those perceiving the sender to be an out-group member only (Study 2).

The fourth and final project aimed at providing a systematic review of political ideology measurement with a special focus on replicability and validity. Using a forward and backward snowballing search strategy, we identified 394 articles of which 207 met all inclusion criteria. Overall, we cataloged more than 60 unique ideological measures, of which only a third had been developed and validated beforehand. About 50% of all identified ideological instruments lacked a single mention of validation evidence. Indeed, the majority of the scales were on-the-fly measures, or a combination of items used in previous studies. Furthermore, the data suggests that replicability might be restricted due to incomplete reporting of the items used, and substantial variance in scoring and scale type. In summary, these circumstances could hinder the ability to build on each other's work and thus likely pose a serious threat to the comparability and generalizability of findings.

The overarching goal of this dissertation was to evaluate the stability of political attitudes and their psychological underpinnings in the face of situational and methodological variance. It highlights the necessity of representative sampling and context sensitivity when assessing the psychological foundations of political attitudes. This becomes even clearer if one takes into account that political ideology by definition develops in shared, that is, social contexts. Acknowledging contextual and methodological variance in the study of political psychology can help us build a more accurate and detailed picture of humans as political beings.

Further information about the PhD-Thesis is available at: https://doi.org/10.22032/dbt.50264

References

Aristotle. (trans. 2009). *Politics*. Oxford World's Classics.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.

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

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

Carmines, E. G., & Stimson, J. A. (1980). The two faces of issue voting. *The American Political Science Review*, *74*(1), 78–91. https://doi.org/10.2307/1955648

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, *130*(6), 959–988. https://doi.org/10.1037/0033-2909.130.6.959

Eckstein, K. (2019). Politische Entwicklung im Jugend- und jungen Erwachsenenalter. In B. Kracke & P. Noack (Eds.), *Handbuch Entwicklungs- und Erziehungspsychologie* (pp. 405–423). Springer. https://doi.org/10.1007/978-3-642-54061-5

James, W. (1975). *The meaning of truth.* Harvard University Press.

Jennings, M. K. (1990). The crystallization of orientations. In M. K. Jennings et al. (Eds.), *Continuities in political action: A longitudinal study of political orientations in three western democracies* (Vol. 5, pp. 313–348). W. de Gruyter.

Kessler, T., Proch, J., Hechler, S., & Nägler, L. (2015). Political diversity versus stimuli diversity: Alternative ways to improve social psychological science. *Behavioral and Brain Sciences*, *38*, E148. https://doi.org/10.1017/S0140525X14001241

Piepenbrink, K. (2001). *Politische Ordnungskonzeptionen in der attischen Demokratie des vierten Jahrhunderts v. Chr: Eine vergleichende Untersuchung zum philosophischen und rhetorischen Diskurs*. Franz Steiner Verlag.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (Vol. 56, p. 33–47). Brooks/Cole.

A PhD Journey in the Steps of Brunswik and Hammond

Darren Conlon

University of Sydney, Australia

University of Notre Dame, Australia

✉ dcon8955@uni.sydney.edu.au; darren.conlon@nd.edu.au

I was first introduced to the work of Egon Brunswik during the research phase of my current PhD journey, where I examined decision-making processes of nurses working in mental health pertaining to confidentiality and disclosure in the context of patient risk. There is a public interest in confidentiality, but there is a competing public interest in releasing confidential information to protect a patient or others from harm. Decision-making in this area of practice is complex because withholding or disclosing information inappropriately creates risks for stakeholders (Conlon et al., 2019; 2021). This study focused on understanding the cognitive processes underpinning the fine balancing act nurses must undertake to make these decisions.

One of the core components of Brunswik's lens theory is that decision-makers are often faced with complex unknown environments they must perceive accurately to act effectively (Brunswik, 1956; Dhami & Mumpower, 2018; Goldstein, 2004). Brunswik proposes that each environment consists of overt and covert areas, and one such covert area is patient risk (Brunswik, 1956; Conlon et al., 2022). Therefore, in assessing patient risk a decision-maker must analyse data-laden proximal cues to discern risk-related distal variables of the environment (Brunswik, 1956; Goldstein, 2004). Furthermore, building on Brunswik's work, Kenneth Hammond's Cognitive Continuum Theory (CCT) holds that decisions may be made by a blending or oscillating between intuitive and analytical approaches (Bjørk & Hamilton, 2011; De Neys, 2022; Dhami & Thompson, 2012; Hamm, 1988; Hammond, 1996).

Principles outlined by both Brunswik and Hammond are analogous with the decision-making approach to clinical risk known as Structured Clinical Judgement (SCJ), whereby nurses' (and indeed all clinicians') intuition is applied to analytical (actuarial) indicators of risk to create a dynamic approach to risk assessment and management (Conlon et al., 2019). The analytical approach in this instance is informed by rules of confidentiality before a final decision is made (Conlon et al., 2022). Furthermore, SCJ recognises the importance of the decision-maker as an instrument of the decision-making process, whereby intuition is an important arbitrator in difficult and sometimes incongruent scenarios (Conlon et al., 2022; Dhami & Mumpower, 2018; Goldstein, 2004). Therefore, Brunswik and Hammond have created a theoretical framework that aligns with my PhD research examining clinical decision-making of nurses working in mental health, regarding confidentiality and disclosure in the context of risk (Conlon et al., 2022).

Author's note: This PhD project takes place at the University of Sydney (Australia) under the mentorship of Associate Professor Timothy Wand and Associate Professor Toby Raeburn.

References

Bjørk, I. T., & Hamilton, G. A. (2011). Clinical decision making of nurses working in hospital settings. *Nursing Research and Practice*, *2011*, 1–8. https://doi.org/10.1155/2011/524918

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

Conlon, D., Raeburn, T., & Wand, T. (2019). Disclosure of confidential information by mental health nurses, of patients they assess to be a risk of harm to self or others: An integrative review. *International Journal of Mental Health Nursing*, *28*(6), 1235–1247. https://doi.org/10.1111/inm.12642

Conlon, D., Raeburn, T., & Wand, T. (2021). Decision-making processes of a nurse working in mental health, regarding disclosure of confidential personal health information of a patient assessed as posing a risk. *Collegian*, *28*(3), 261–267. https://doi.org/10.1016/j.colegn.2020.08.010

Conlon, D., Raeburn, T., & Wand, T. (2022). Cognitive Continuum Theory: Can it contribute to the examination of confidentiality and risk-actuated disclosure decisions of nurses practicing in mental health? *Nursing Inquiry,* e12520. https://doi.org/10.1111/nin.12520

De Neys, W. (2022). Advancing theorizing about fast-and-slow thinking. *The Behavioral and Brain Sciences*, 1–68. https://doi.org/10.1017/S0140525X2200142X

Dhami, M. K., & Mumpower, J. L. (2018). Kenneth R. Hammond's contributions to the study of judgment and decision making. *Judgment and Decision Making*, *13*(1), 1–22. http://journal.sjdm.org/vol13.1.html

Dhami, M. K., & Thomson, M. E. (2012). On the relevance of cognitive continuum theory and quasirationality for understanding management judgment and decision making. *European Management Journal*, *30*(4), 316–326. https://doi.org/10.1016/j.emj.2012.02.002

Goldstein, W. M. (2004). Social Judgment Theory: Applying and extending Brunswik's Probabilistic Functionalism. In D. J. Kohler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 37–61). Blackwell Publishing.

Hamm, R. M. (1988). Moment-by-moment variation in experts' analytic and intuitive cognitive activity. *IEEE Transactions on Systems, Man, and Cybernetics*, *18*(5), 757–776. https://doi.org/10.1109/21.21602

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

Coherence and Correspondence Reconceptualized

Michael E. Doherty
Bowling Green State University, United States
✉ mdohertyjdm@gmail.com

R. James Holzworth
University of Connecticut, United States
✉ jimholzworth@gmail.com

Thomas R. Stewart
University at Albany, United States
✉ t.stewart@albany.edu

Ken Hammond (1996, p. 95) wrote: *Two grand metatheories have been persistent rivals in the history of science in general and in the history of research in judgment and decision making.*

When he applied the classic distinction between correspondence and coherence theories of truth to the field of judgment and decision research (Hammond, 1996, chapter 4), we thought it was inspired. In one stroke he clarified a major difference between heuristics and biases studies and lens model studies and helped explain why they relied on different paradigms and reached different conclusions. …" The basic idea of correspondence theory is that what we believed is true if it corresponds to the way things actually are. That is, the truth of a proposition is established by its relationship to the world. In coherence theory, the truth of a proposition is established by its relationships to other propositions. We believe that coherence is often instrumental in achieving correspondence.

Now, after more than 40 years, we believe that it is time to reexamine, refine and build on Hammond's insights. We build on that foundation laid by Hammond when we reconceptualize the two theories of truth as continua rather than simpler dichotomies. We label the added continua the *Chaos–Coherence* and *Irrelevance–Correspondence* continua.

The term "metatheory" is not typical in conversations in psychology laboratories. It does not denote anything like a theory in the sense of a testable proposition about reality. It is not falsifiable. It refers not to a theory about some aspect of the natural world. It refers to theory about a theory. Metatheories are crucial as guides to theoretical progress. They may never be subjected to questions like whether they can meet a $p < .01$ level, but they can be widely accepted – or not – by the scientific community as criteria for theory. In the present context, metatheory of truth refers to a framework by which we assess truth of a theory of judgment and decision making.

Our proposed continua are part of *Cognitive Continuum Theory* (CCT) that evolved through the 1980s and 1990s (Hammond, 1980, 1981, 1986, 1987, 1988, 1990a,b) and continue to evolve.

CCT is detailed in the book by Hammond (1996) that was given an award by the American Educational Research Association (AERA) for its contribution to theory and research in education.

The metatheories of truth considered as cognitive continua

In the above (opening) quotation Hammond identified an important feature of the history of science. Both small advances in scientific knowledge and scientific revolutions involved interpretations of theories and observations in light of the metatheories of truth, *coherence* and *correspondence.* Theories and data inexorably change, but the idea of the truth criteria of coherence and correspondence endure. We see truth metatheories not only as important in the history of science, but also as representing what is happening in the minds of scientists. The quote from Hammond may not have been intended to describe the day-to-day behavior of the working scientist, but we propose that it is directly and deeply relevant thereto. When scientists think of designing an experiment, we speculate that coherence thinking marks the design and analysis of the whole experiment, whereas we speculate that correspondence thinking marks especially the thinking of and selection of the dependent variable. But, in one way or another, every aspect of research and theory are evaluated in light of the metatheories. These speculations certainly are not based on "hard" data, but on the lived experience of scientists relevant to understanding and analyzing the construction of theory and the design of experiments. Hammond's conception that intuition and analysis lie on a continuum resonates with our lived experience! As will be clear below, we argue that coherence and correspondence are not fixed, immutable criteria but rather two modes of thought on continua that are in some ways analogous to Hammond's conceptualization of analysis and intuition as end points of cognitive continua. We note at the outset that considerable theoretical work will be necessary to flesh out the details of the two additional cognitive continua.

The *Intuition–Analysis* continuum: As noted above Hammond argued that intuition and analysis should be treated as the end points of a cognitive continuum rather than a dichotomy. The continuum framework embodies psychological coherence much more faithfully than does the dichotomous framework. It is a metatheory that has not and cannot be empirically tested but it has been widely accepted by the scientific community of which we are part. We could explore these considerations further, but we'd be preaching to the choir. Let us say *Kudos to Ken* for CCT and move on.

The *Irrelevance–Correspondence* continuum: We propose that scientists check their understanding of a proposition that is on the *Intuition-Analysis* cognitive continuum against one or both of the truth continua. That is, the scientist has internalized a set of principles concerning expectations of outcomes. These expectations may be based on similar studies in the relevant literature, pilot studies etc. The conditions of the research and expectations of results are then matched to what is on the scientist's *Irrelevance-Correspondence continuum*.

At the extreme irrelevance end of the continuum, we easily dismiss scientific claims that we believe violate nature. The proponent may see the predicted outcome of a planned study as almost universal – expecting a pretty high level of correspondence – but the scientific community or representatives thereof – sees proposition as completely lacking any predictive or explanatory power. The other extreme end of the continuum involves a situation in which the evidence is overwhelming: The predictions match reality as measured. The focus of interest in scientific

thinking lies between the two ends of the continuum. It is there that the work of science is done and that science progresses.

We believe that it is almost self-evident that correspondence is a matter of degree. Correspondence in the social judgment theoretical tradition is assessed by the accuracy coefficient, r_{YE} which is intrinsically continuous. In a Bayesian study in the tradition of Ward Edwards the criterion for success of a model is typically a conditional probability, which is intrinsically continuous. The very notion that science progresses over time means that at time 1 the scientist checks understanding and concludes that the criterion for truth has been met, then perhaps at times 2, 3, etc. new kinds of information surface. Revision of understanding on *the Intuition–Analysis* continuum occurs. The changed position is checked against the truth criterion. New kinds of information may require change on the *Irrelevance–Correspondence continuum* and the work goes on. The concept of continua leaves open the idea that a potential solution may be good or very good yet not perfect. The concept of continua is consistent with the idea of knowledge being tentative.

The Chaos–Coherence continuum: As above, we propose that scientists check their understanding of a proposition that is on the *Intuition-Analysis continuum* against The *Chaos–Coherence continuum*. That is, the scientist has internalized a set of principles defining or describing the concept of coherence. Today's scientist takes advantage of millennia of progress not only because of the vast base of relevant knowledge and technology now available, but also because of the very modes of thought that we call scientific thinking, including the metatheories of truth.

Today's scientist does not start out at the chaos end of the truth continuum, even if the problem is brand new. In some fields a mathematical model may already be available to be adapted or exploited. It is not news to any reader of the newsletter that scientists often struggle for a considerable time to come up with coherent explanations as well as good experiments. Potential solutions may be assessed first by comparison with the *Chaos–Coherence* continuum. At the extreme irrelevance end of the continuum, easily dismissed are scientific claims that rest on propositions that are literally incoherent. The *Chaos–Coherence* continuum stretches all the way to a mathematical or logical model. We have been talking as though science is a personal endeavor. It is, and intensely so. It is also a social endeavor and intensely so. Everything we have just located in the past and in the heads of the individual scientist is clearly in the scientific culture and interactions between scientists. The creative scientist is not the only one assessing theories, experimental outcomes and truth criteria.

The above is clearly a work in progress. We plan to develop these ideas further in a more complete paper, elaborating upon each continuum. We accept the *Intuition–Analysis* continuum as the foundation of *Cognitive Continuum Theory*. The proposed idea of an *Irrelevance–Correspondence* continuum concerns goodness or accuracy of a potential solution. The proposed idea of a *Chaos–Coherence* continuum concerns logic and understanding of a potential solution. Hopefully, we will develop a coherent combination of these metatheories. Please feel free to send suggestions, criticisms or comments or just friendly notes to any one of the authors.

References

Hammond, K. R. (1980). *The integration of research in judgment and decision theory* (Report 226). Center for Research on Judgment and Policy, University of Colorado, Boulder, CO.

Hammond, K. R. (1981). *Principles of organization in intuitive and analytical cognition* (Report 231). Center for Research on Judgment and Policy, University of Colorado, Boulder, CO.

Hammond, K. R. (1986). *A theoretically based review of theory and research in judgment and decision making* (Report 260). Center for Research on Judgment and Policy, University of Colorado, Boulder, CO.

Hammond, K. R. (1987). Toward a unified approach to the study of expert judgment. In J. Mumpower, L. Phillips, O. Renn, & R. Uppuluri (Eds.), *Expert judgment and expert systems*. Berlin: Springer-Verlag, 1–16.

Hammond, K. R. (1988). Judgment and decision making in dynamic tasks. *Information and Decision Technologies*, *14*, 3–14.

Hammond, K. R. (1990a). Functionalism and illusionism: Can integration be usefully achieved? In R. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn*. Chicago: University of Chicago Press.

Hammond, K. R. (1990b). Intuitive and analytical cognition: Information models. In A. Sage (Ed.), *Concise encyclopedia of information processing in systems and organizations* (pp. 306–312). Oxford: Pergamon Press.

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

The Other Side of Brunswik's Organism-Environment Model (Butterfly)

Robert M. Hamm

University of Oklahoma Health Sciences Center, United States

✉ robert-hamm@ouhsc.edu

This figurative representation of Brunswik's organism-environment model (which we could call the "butterfly model", in contrast with the lens model), from Bernhard Wolf's (2005) summary of Brunswik's theory, shows perceptions converging (left wing) from remote to distal to proximal to the organism's periphery and center, and then diverging again (right wing) as actions, from proximal to distal to remote.

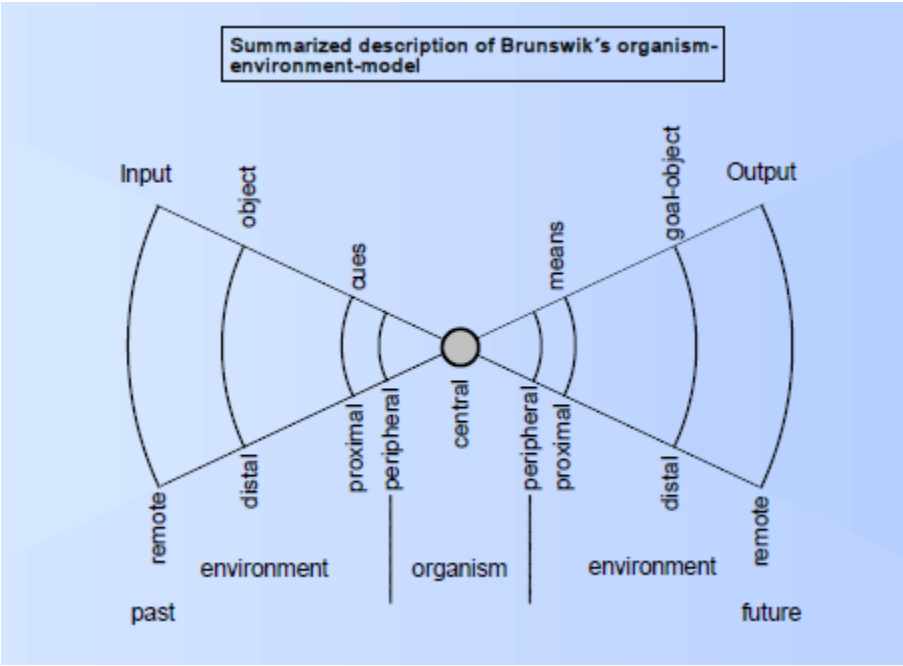


Figure 1. Brunswik's organism-environment model (butterfly graph).

I realized that in using the lens model, I have always focused on the left wing. That is, I and most others (see Kaufmann, Reips, & Wittman, 2013) use the lens to represent cue utilization in perception or prediction, how the central representation of the world relates to the actual past or future environment, independent of our own actions. Cues related to the (class of) remote objects or events are collected, used, integrated into a judgment or perception. But there could be an analogous lens model on the right wing, relating the intended consequences of one's actions to their actual effects, mediated by various aspects of our execution of the action.

The action lens model would have this familiar form, with the subjective on the left rather than the right (to preserve its orientation in the butterfly model).

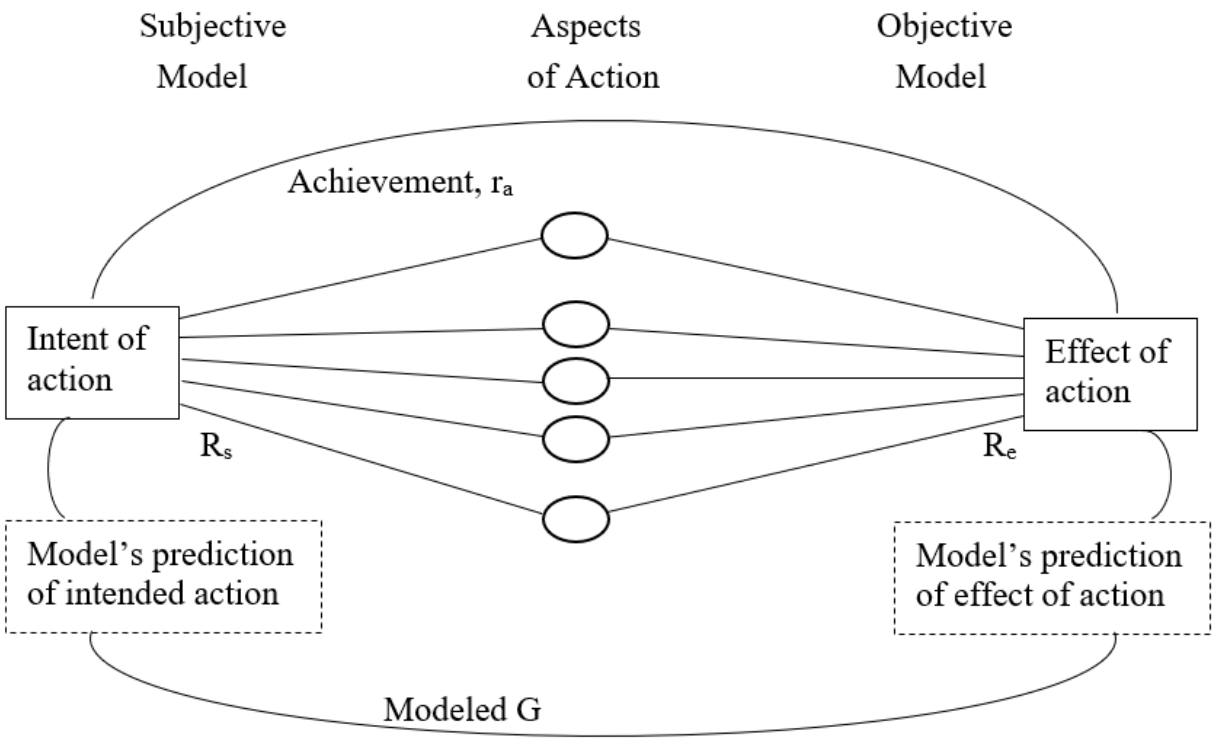


Figure 2. Action lens model.

What might data for an action lens model look like?

Suppose one is playing a beanbag toss game, in which the task is to distribute one's 10 beanbags uniformly over the range 5 meters to 15 meters. Imagine we hide the toss from the player immediately after they release the bag, or even have them throw blindfolded, and ask for a prediction of how far it went. The player also would report features of the toss (timing of release "too early to too late," force "too weak to too strong," aim "too high to too low"). Or the tosses could be filmed and objective measures of the release and beginning of the arc also derived.

For analysis, "achievement" would ignore the rules of this game and focus instead on the accuracy of the prediction. The predicted distance would be correlated with the actual distance, for r_a. For models of the intended action (distance), and of the action's actual effect, the distances would be regressed onto the aspects of the action. The predictions and residuals would be used in a lens model, and assessment of the causes of inaccurate judgment could compare the relative cue weights in the regressions of the predicted distances and the actual distances. Given that "there's many a slip 'twixt the cup and the lip," perhaps it would be fruitful to decompose the connection between intended action and actually performed action, or other aspects of the causal connections, analogous to the approach described by Stewart and Lusk (1994).

Possibly someone has already made lens models of actions on the world. Araújo and Kirlik (2008) approach it, discussing the lens model in relation to visual anticipation in sport, but their focus was more on predicting other players' moves and ball movements, rather than predicting the effects of one's own actions. The same holds in seven papers reviewed in Kaufmann et al. (2013). (Let us

know of any examples you are aware of.) Reading an earlier draft of this, Len Adelman and Tom Stewart each thought of the US Federal Reserve Bank's directors, trying to control the rate of inflation (intent of action) by adjusting the prime interest rate, the purchase of bonds, and other aspects of the economy that they might have control over (aspects of actions). For psychology teachers, researchers seeking projects, and coaches of skilled performance, it is useful to recognize this possible area of application of Brunswik's theory.

References

Araújo, D., & Kirlik, A. (2008). Towards an ecological approach to visual anticipation for expert performance in sport. *International Journal of Sport Psychology*, *39*(2), 157–165.

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), e83528. https://doi.org/10.1371/journal.pone.0083528

Stewart, T. R., & Lusk, C. M. (1994). Seven components of judgmental forecasting skill: Implications for research and the improvement of forecasts. *Journal of Forecasting*, *13*(7), 579–599. https://doi.org/10.1002/for.3980130703

Wolf, B. (2000). *The Structure of the Human World: Brunswik's Organism-environment-model*. Retrieved November 22, 2022, from https://brunswiksociety.org/wp-content/uploads/2022/06/essay6.pdf.

The Potential of Digitalization and Advanced Analysis Techniques: Lens Model Studies for Teacher Education and Training

Esther Kaufmann

University of Konstanz, Germany

✉ esther.kaufmann@uni-konstanz.de

Lens model studies focus mostly on research and are less used for education and training. However, in line with Wigton (2008), lens model studies also have great potential not only for medical education but, I argue, also for teacher education and training.

In my recently finished and successfully defended habilitation thesis at the University of Zurich (Switzerland), I emphasized the potential of the lens model approach for education science. In one subchapter, I showed that the current development of science and digitalization potentially has an impact on advancing the application of lens model studies, which may have been underestimated until now. More precisely, I argue that digital archiving of individual data and their analysis with advanced techniques (i.e., individual participant data meta-analysis) increases the potential of lens model studies for teacher education and training.

In the following, I provide a short overview on my ideas. First, I introduce an example of a classical lens model study. Cooksey et al. (1986) is an excellent example of applying a lens model study to teachers' judgment achievement (see Figure 1). In this study, 20 teachers were asked to evaluate the reading comprehension of 118 children (see the right side of Figure 1, Y_s). Their evaluations were based on written profiles which included five important cues (e.g., information about each child's reading ability, oral language ability, and socioeconomic status; see Figure 1, $X_{1\text{-}k}$). Teachers' judgments were then compared with students' scores on a reading comprehension test (left side of Figure 1, Y_e), a criterion that has also been used in non-lens model studies (see Kaufmann, 2020; Urhahne & Wijnia, 2021). The correlation between teachers' judgments and the student's test scores was the indicator of judgment achievement, such that a higher correlation indicated more accurate judgment. Correlational statistics were also used to capture the relationships between cues and teachers' judgments (cue utilization, Figure 1), cues and the environmental criterion, reading comprehensive test (ecological validity, see Figure 1), as well as inter-cue correlations (for details, see lens model equation, Tucker, 1964).

In Kaufmann (2022), an overview of lens model studies within the education field is also available and, ideally, supplements research on teachers' judgment accuracy in non-lens model studies (see Kaufmann, 2020; Urhahne & Wijnia, 2021). However, I was only able to find five lens model studies considering 93 individual judgments, showing a missing potential of lens model studies within the educational field. Because this database, consisting of several pieces of information about individuals, is of great importance for teachers' development and education. I suggest using this database as a benchmark and adding additional data to it by running lens model studies with teachers. The incoming data should integrate into this database by a so-called Individual Participant Data (IPD)-meta-analysis to prevent any aggregation bias. With such an approach, data from individual teachers could be integrated into the previous database and compared with

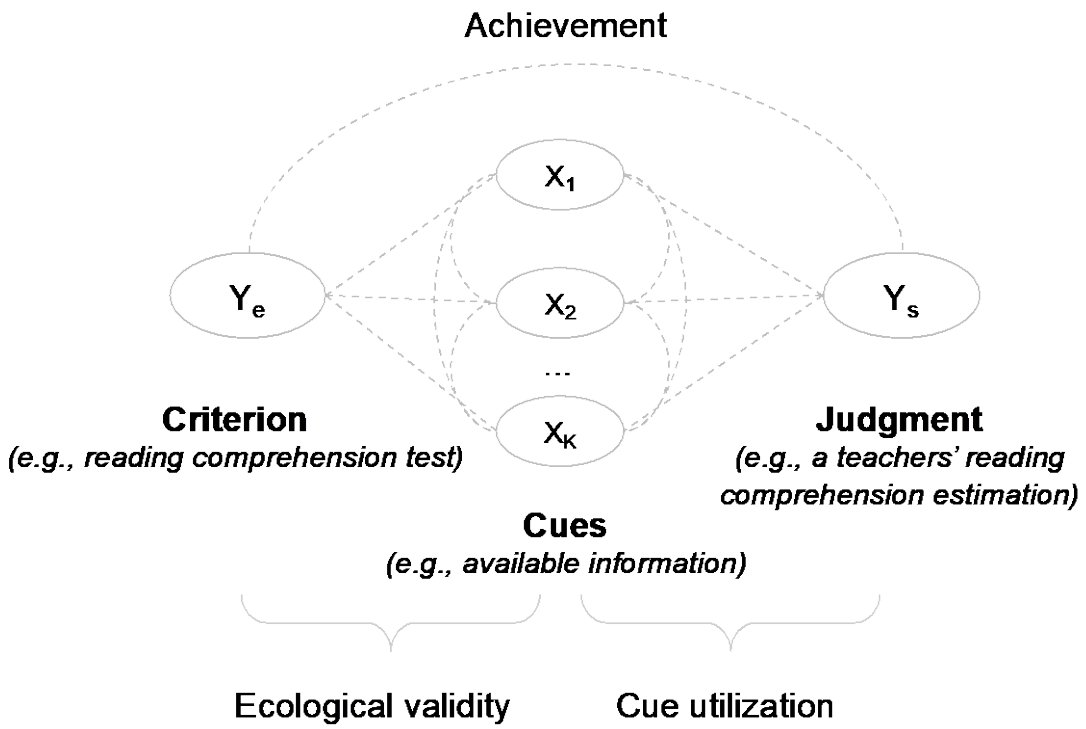


Figure 1. Classical lens model.

previous teacher data on the same task. This comparison is ideal to provide feedback to teachers about their judgment accuracy. The current increase in digitalization reduces the workload of such an undertaking because data storage and running an IPD meta-analysis is getting more feasible than in years before. Therefore, I see the potential in the increased digitalization combined with the lens model approach to improve teachers' judgment accuracy and finally, possibly to reduce any injustice between students by inaccurate teachers' judgments.

Further information about my suggestion is given by Kaufmann (2022). Feedback is welcomed and greatly appreciated.

References

Kaufmann, E. (2022). Lens model studies: Revealing teachers' judgements for teacher education. *Journal of Education for Teaching (*JET*)*. https://doi.org/10.1080/02607476.2022.2061336

Kaufmann, E. (2020). How accurately do teachers' judge students? Re-analysis of Hoge and Coladarci's (1989) meta-analysis. *Contemporary Educational Psychology*, *63*, 101902. https://doi.org/10.1016/j.cedpsych.2020.101902

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

Urhahne, D., & Wijnia, L. (2021). A review on the accuracy of teacher judgments. *Educational Research Review*, *32*, 100374. https://doi.org/10.1016/j.edurev.2020.100374.

Wigton, R. S. (2008). What do the theories of Egon Brunswik have to say to medical education? *Advance in Health Science Education*, *13*, 109–121. https://doi.org/10.1007/s10459-006-9023-5

Extracting the Wisdom of a Smaller Crowd from Dependent Quantile Judgments

Yuanyuan Lei & Chen Wang

Tsinghua University, China

✉ chenwang@tsinghua.edu.cn

The wisdom of a crowd arises from the combining of diversified judgments. There is extensive literature on the aggregation of quantile judgments. The simplest and most common approach is simple averaging (Clemen, 1989; Lichtendahl, Grushka-Cockayne, & Pfeifer, 2013; Makridakis & Winkler, 1983). Given historical data about experts' past performance, performance-based aggregation can achieve the desired predictive accuracy in theory (Bansal et al., 2017; Budescu & Chen, 2015; Cooke, 1991; Hora, 2004; Winkler & Clemen, 1992). However, the predictability and interpretability of expert judgments are both essential for aggregation. One explanation of dependence between experts is the "shared-information problem" (Palley & Soll, 2019).

Therefore, information extraction is especially beneficial when the true outcome is far in the future (e.g., long-term consequences of a public policy, inflation in years to come, etc.), and in the case of peer predictions where no outcome can be verified despite mechanisms to encourage truth-telling (Miller et al., 2005). The lens model is an excellent choice to model how an expert processes multiple pieces of information, with both theoretical ground and analytic convenience.

The lens model explains how people perceive the environment they live (Brunswik, 1939). Understanding of a variable of interest is not obtained directly from the environment but inferred from some information cues (Wolf, 2000). Following this idea, we propose a judgment model where each expert is assumed to combine multiple information cues linearly, subject to prediction errors in their quantile judgments (Lei & Wang, 2022).

Suppose that a decision-maker consults K experts to make predictions for I variables of interest. Denote $Y_{ki}(t)$ as expert k's quantile judgment about variable i at probability $t \in (0,1)$ for $k = 1, \cdots, K$ and $i = 1, \cdots, I$. Denote $\phi_{qi}(t)$ as the quantile function of the qth information cue with $\beta_{kq} \geq 0$ and $\sum_{q=1}^{Q} \beta_{kq} = 1$. We assume that each expert obtains the well-calibrated quantile value $X_{ki}(t)$ by linearly combining Q information cues. Specifically, expert k's quantile judgment can be given by $Y_{ki}(t) = \sum_{q=1}^{Q} \beta_{kq} \phi_{qi}(t) + \varepsilon_{ki}(t)$, where $\varepsilon_{ki}(t)$ is prediction error consisting of sampling and judgmental errors. We use the Gaussian process to account for correlations of sampling and judgmental errors between quantile judgments at different probability levels.

By utilizing the lens model to understand expert dependence, we can separate the sampling and judgmental errors from the well-calibrated quantile values and pinpoint each expert's information, reflecting the underlying diversity and dependence. Inspired by matrix factorization and the lens model, we simultaneously extract "variable profiles" (latent cues that underlie the quantile judgments of multiple variables) and "expert profiles" (how much weights each expert places on these cues) from a given set of quantile judgments. Based on the linear information structure, clustering experts by their weights over the latent cues is equivalent to clustering by weights over

the actual cues (fully or partially unobservable). Moreover, the lens model supports the validity of averaging quantiles to aggregate expert predictions.

Dependence among expert judgments reduces the merit of each forecast (Clemen & Winkler, 1985; Ferrell & Gresham, 1985; Hogarth, 1978; Wallsten et al., 1997). We then show by simulation and case studies that clustering experts according to their estimated weights and picking one delegate from each group are sufficient to represent the entire panel. Selecting a subset of experts can reduce organizational costs and combat overfitting in a new elicitation task. For example, we analyzed the quantile forecasts of annual GDP growth rates from the Federal Reserve Bank of Philadelphia (FED)'s Surveys of Professional Forecasters (SPF). Eight forecasters selected by our method were representative of twenty forecasters. One of the estimated latent cues also highly correlated with the real GDP growth rate, indicating that the FED's experts indeed had access to reliable evidence about economic growth.

References

Bansal, S., Gutierrez, G. J., & Keiser, J. R. (2017). Using experts' noisy quantile judgments to quantify risks: Theory and application to agribusiness. *Operations Research*, *65*(5), 1115–1130.

Brunswik, E. (1939). The conceptual focus of some psychological systems. *The Journal of Unified Science (Erkenntnis)*, *8*(1), 36–49.

Budescu, D. V., & Chen, E. (2015). Identifying expertise to extract the wisdom of crowds. *Management Science*, *61*(2), 267–280.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, *5*(4), 559–583.

Clemen, R. T., &Winkler, R. L. (1985). Limits for the precision and value of information from dependent sources. *Operations Research*, *33*(2), 427–442.

Cooke, R. M. (1991). *Experts in uncertainty: Opinion and subjective probability in science*. Oxford University Press, New York, NY.

Ferrell, O. C., & Gresham, L. G. (1985). A contingency framework for understanding ethical decision making in marketing. *Journal of Marketing*, *49*(3), 87–96.

Hogarth, R. M. (1978). A note on aggregating opinions. *Organizational Behavior and Human Performance*, *21*(1), 40–46.

Hora, S. C. (2004). Probability judgments for continuous quantities: Linear combinations and calibration. *Management Science*, *50*(5), 597–604.

Lei, Y., & Wang, C. (2022). Extracting the wisdom of a smaller crowd from dependent quantile judgments. *IISE Transactions*, https://doi.org/10.1080/24725854.2022.2086326.

Lichtendahl K. C., Jr, Grushka-Cockayne, Y., & Winkler, R. L. (2013). Is it better to average probabilities or quantiles? *Management Science*, *59*(7), 1594–1611.

Makridakis, S., & Winkler, R. L. (1983). Averages of forecasts: Some empirical results. *Management Science*, *29*(9), 987–996.

Miller, N., Resnick, P., & Zeckhauser, R. (2005). Eliciting informative feedback: The peer prediction method. *Management Science*, *51*(9),1359–1373.

Palley, A. B., & Soll, J. B. (2019). Extracting the wisdom of crowds when information is shared. *Management Science*, *65*(5), 1–19.

Wallsten, T. S., Budescu, D. V., Erev, I., & Diederich, A. (1997). Evaluating and combining subjective probability estimates. *Journal of Behavioral Decision Making*, *10*(3), 243–268.

Winkler, R. L., & Clemen, R. T. (1992). Sensitivity of weights in combining forecasts. *Operations Research*, *40*(3), 609–614.

Wolf, B. (2000). *The Structure of the Human World: Brunswik's Organism-environment-model*. Retrieved March 12, 2021, from http://www.brunswik.org/.

Using the Extended Lens Model to Reflect on the Ethical Use of Artificial Intelligence

Luis F. Luna-Reyes & Teresa M. Harrison

University at Albany, United States

✉ lluna-reyes@albany.edu ✉ tharrison@albany.edu

"Insight – Using data, evidence and analytics to create insight that influences decision making, actions and results" has been identified as a major driver for government transformation (Chenok et al., 2017). Government's vast stores of data can now be computationally digested with the use of sophisticated analytical and cognitive tools to produce "actionable" insight –which improves decision making, produces efficiencies, and stimulates innovation across the wide range of government operations. In this turn toward data-driven analysis and decision making, government seeks to exploit the "disruptive" technologies of Artificial Intelligence (AI), an umbrella term that refers to a set of sophisticated computational strategies, including machine learning, neural networks, natural language processing, autonomous vehicles, computer vision, and facial recognition. These strategies promise competitive advantages to various industries as well as cost savings and innovations for government (e.g., Eggers et al., 2017).

While widely regarded as revolutionary, the benefits and costs of deploying AI strategies in government contexts are still emerging. Industry analysts have recognized that AI may challenge democratic institutions in significant ways (e.g., Araya, 2019), also producing substantial threats to privacy, autonomy, equity, and fairness for individuals in their quotidian pursuits and in their roles as citizens (e.g., Eubanks, 2018). This deep integration of AI driven information technology in almost every aspect of public policy and public service development threatens to alter the practice and experience of governance in ways that are only beginning to be appreciated.

In our recent paper, "Cultivating Trustworthy Artificial Intelligence in Digital Government" (Harrison & Luna-Reyes, 2022), we reflect on the role of multi-criteria decision making, and the Lens Model in particular (Dhami & Mumpower, 2018), in responding to major questions in this transformation process and developing trustworthy AI applications. Specifically, we use the Extended Lens Model (ELM) proposed by Stewart and Lusk (see Figure 1) to think about major challenges on the development of AI systems.

We identified four major challenges through the ELM. *The fidelity of the information system* includes challenges emerging from the integration of multiple sources of information assembled in contexts with different and potentially unknown data management practices. Such integration – necessary for all AI applications—must be complete, valid, and accurate. In practice, high variability in the validity and quality of the data is pervasive. Although AI systems tend to be perceived as objective, many different judges are needed to *reliably acquire information* during the processes of data cleaning and integration. Most of these judgments are likely made by software engineers with little or no input from the decision makers, and with poor documentation practices. Increased openness in the process, as well as improved understanding of the impact of such judgments in the system are necessary and pose challenges for training and collaboration between engineers and public managers. The ELM calls for *robust environmental models,* which

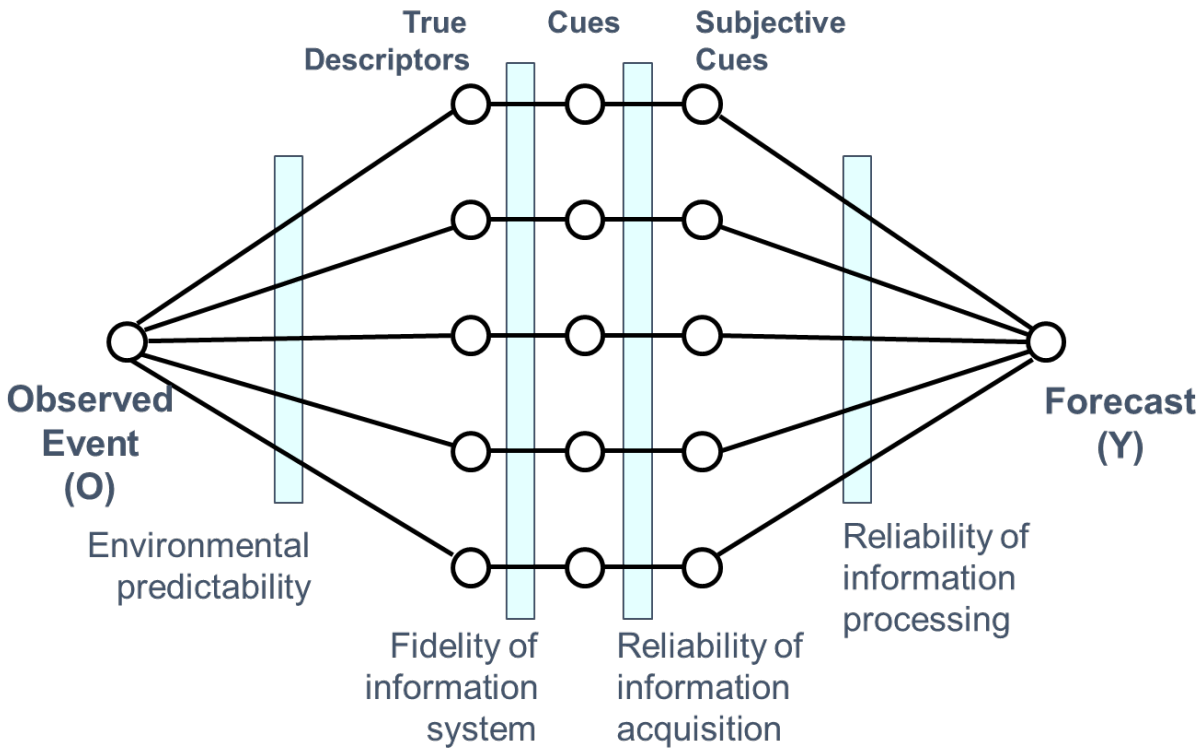


Figure 1. Extended Lens Model (adapted from Stewart & Lusk, 1994).

traditionally relies on domain-based model building, which poses a challenge to many AI strategies (e.g., deep neural networks) that often eschew both theory and domain knowledge focusing instead on using available data to locate a combination of predictive factors and derive a set of decision rules. Finally, best practices on multi-criteria decision approaches includes the *assessment of the forecast*, interrogating the validity of fundamental assumptions and looking for the adequacy of the forecast. In contrast, AI strategies are not necessarily amenable to human assessment and critique. This is because characteristics of learning algorithms make it difficult to examine internal decision processes. The complexity of model development makes it difficult for even experts to fully grasp the nature of the calculations, with significant implications for explaining how an algorithm works or representing its logic to those affected by the decisions made.

These four issues relate to two major challenges of data and algorithm requirements to develop trustworthy AI systems. The first one relates to improving data quality by promoting more integrated data management practices across government agencies and increased data literacy among government employees so that they can actively collaborate in producing better curated datasets as well as protecting the privacy and data sovereignty of citizens. The second strategy involves the development of AI governance strategies that include the widest range of stakeholders including policy makers, government domain experts, AI systems developers inside and external to government, along with individuals who will be affected by AI decision making. Actively involving all these actors in the planning, development, assessment and audits of AI strategies and applications is necessary for the deployment of trustworthy AI. Clearly, we need to know more about the ways in which AI and its associated data are being used by government at all levels.

Using established models and approaches to decision making are promising guides to produce better and more just systems.

References

Araya, D. (2019, April 1). Artificial intelligence and the end of government [Blog]. *Forbes*. https://www.forbes.com/sites/danielaraya/2019/01/04/artificial-intelligence-and-the-end-of-government/#1c70a13b719b

Chenok, D. J., Cooney, H., Kamensky, J. M., Keegan, M. J., & Piechowski, D. (2017). *Seven drivers transforming government* (Special Report Series, p. 46). IBM Center for the Business of Government. http://www.businessofgovernment.org/sites/default/files/Seven%20Drivers%20Transforming%20Government.pdf

Dhami, M. K., & Mumpower, J. L. (2018). Kenneth R. Hammond's contributions to the study of judgment and decision making. *Judgment and Decision Making*, *13*(1), 1–22.

Eggers, W. D., Schatsky, D., & Viechnicki, P. (2017). *AI-augmented government: Using cognitive technology redesign public sector work*. Deloitte Center for Government Insights.

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.

Harrison, T. M., & Luna-Reyes, L. F. (2022). Cultivating trustworthy artificial intelligence in digital government. *Social Science Computer Review*, *40*(2), 494–511. https://doi.org/10.1177/0894439320980122

Stewart, T. R., & Lusk, C. M. (1994). Seven components of judgmental forecasting skill: Implications for research and the improvement of forecasts. *Journal of Forecasting*, *13*(7), 579–599. https://doi.org/10.1002/for.3980130703

The Accurate Judgment of Social Network Characteristics in the Lab and Field Using Thin Slices of the Behavioral Stream

Sanaz Mobasseri

Boston University, United States

✉ sanazm@bu.edu

Daniel H. Stein

University of California, Berkeley, United States

Dana R. Carney

University of California, Berkeley, United States

Humans demonstrate a remarkable, perhaps biological, readiness to form and maintain memberships in groups and construct social networks (Baumeister & Leary, 1995; Durkheim, 1995). In fact, our long-term survival necessitates the successful formation and maintenance of group membership and social networks (Dunbar, 2008). Specifically, membership in social groups and networks allows us to rely on others for information, protection, aid, and resources, which maximizes our likelihood of survival (Brewer, 1997; Caporael, 1997; Ibarra et al., 2005; Lin, 2001). At the same time, affiliating with a stranger brings potential risk to the self and one's network. Therefore, when we first encounter a stranger, we must decide whether or not to affiliate with them and, if so, to what degree. Thus, one may be afforded advantage from making accurate inferences about others' states of mind, intentions, personality, and the company they keep—the latter of which science knows very little about thus, the focus of the current research.

Have humans evolved to be able to accurately assess the quantity and quality of a stranger's relationships with others? Across three lab studies and one preregistered field study, we tested whether people (total $N = 1{,}545$) could make accurate judgments about a stranger's (total $N = 709$) social network characteristics after watching brief thin slice videos of the stranger or negotiating with them. Specifically, we tested whether perceivers could accurately detect four egocentric network characteristics of strangers: size of the target's social network, composition (share of males vs. females and share of family vs. non-family in the target's network), and interconnectedness (how many of the target's social ties know one another) (Bernard et al., 1984; Moreno, 1934; Perry et al., 2018; Rossi, 1966; Wellman, 1993). We also explored how accurate judgments were possible by examining the role of theoretically relevant social-behavioral tendencies and traits (Ickes, 1993).

Study 1 contained two phases: in Phase 1, we constructed videotaped stimuli with rich criterion data on target participants' social networks ($n = 23$). In Phase 2, participant perceivers ($n = 375$) watched short, standardized videos (i.e., thin slices) of these targets and made judgments about their network characteristics, allowing us to compute accuracy coefficients separately for each of the four network characteristics. Study 1 demonstrated that perceivers could make accurate inferences at zero acquaintance about the size ($r = 0.09$, $p < .001$) and composition (gender, $r = 0.33$, $p < .001$; family, $r = 0.07$, $p < .001$) of a target's social network but not about

interconnectedness among target's contacts. Studies 2a and 2b replicated the results from Study 1 about the size ($r = 0.12$, $p < .001$ in Studies 2a and 2b), family composition ($r = 0.18$, $p < .001$ in Study 2a; $r = 0.16$, $p < .001$ in Study 2b), and gender composition ($r = 0.46$, $p < .001$ in Study 2a; $r = 0.50$, $p < .001$ in Study 2b) of a target's social network with two new samples of perceivers ($n = 212$, and $n = 272$). A secondary goal of Studies 2a and 2b was to construct and provide preliminary validation data for an individual difference measure that assesses the degree to which an individual can accurately judge social network characteristics, called the Social Network Accuracy Test (SNAT). The SNAT, a 10-item video test, was made freely available to the research community to test additional questions not asked in the current research. All target videos and data, code, as well as raw data for additional network characteristics not tested in the current research, can be found here: https://osf.io/zgbse. Overall, the evidence from 859 perceivers across Studies 1, 2a, and 2b is the first to demonstrate that perceivers can accurately detect network size, gender, and family composition, but not interconnectedness, after watching thin-slice videos of ordinary people engaging in routinely expressive behaviors for about 100 seconds.

Study 3 extended this work by conceptually replicating our effects using an actual live dyadic negotiation task rather than pre-recorded videos. Study 3 increased the ecological validity of the research and greatly expanded and diversified our sample of perceived targets. Thus, with a larger target sample size (twenty-fold larger, $n = 686$), Study 3 was also able to investigate, with sufficient statistical power, which social behaviors perceivers used correctly to make accurate judgments about social network characteristics. Drawing on Brunswik's (1956) lens model, we conducted a series of analyses to test (a) the extent to which inferences about the target's social behaviors and characteristics were related to what the target reported about their social network characteristics, and (b) the extent to which perceivers used inferences about the target's social behaviors and characteristics to make inferences about the target's social network characteristics. We show that perceivers used behavioral cues about social-behavioral tendencies and traits—for example, related to sociability and gender—to make judgments about the target's social network size ($r = 0.08$, $p = .043$), gender composition ($r = 0.38$, $p < .001$), and family composition ($r = 0.11$, $p = .005$), but not interconnectedness. Cues were sometimes correctly used and sometimes incorrectly used, and at other times, cues that could have been used to facilitate accuracy were ignored. Study 3 results demonstrate that even in an ecologically valid, face-to-face negotiation in which participants were both stressed and cognitively taxed, perceivers were still able to accurately infer the network characteristics of their interaction partner in a manner consistent with the thin-slice lab studies reported in Studies 1, 2a, and 2b. These data reveal that the accuracy effects reported here are replicable, reliable, and generalizable, and that any effects observed using the SNAT can likely be trusted, as they are consistent with data harvested using a completely different, real-world, face-to-face paradigm.

Overall, our findings consistently demonstrated that perceivers accurately detected the size of a stranger's social networks and their gender and family composition, based on theoretically relevant social-behavioral tendencies and traits (e.g., extraversion, gender), but not how interconnected these social networks were. Perceivers also missed cues that could have facilitated greater accuracy. This set of studies is the first of its kind, of which we are aware, to demonstrate people's ability to accurately judge some of the most critical aspects of a stranger's social network. We also provide the freely available Social Network Accuracy Test (SNAT) for future research—(https://osf.io/zgbse)—an individual difference measure that assesses the ability to accurately detect social network characteristics. In so doing, we have made available target videos and

untested social network characteristics (e.g., racial composition) that we hope will enable future researchers to test additional hypotheses about social network accuracy. Future research may use these data—which include targets' self-reported cues (e.g., personality), perceivers' judgments about targets' cues, targets' self-reported social networks, and perceivers' judgments about targets' social networks—to construct several full Brunswik lens models. Such lens models using all cues to decompose accuracy can shed light on numerous questions about group, not only individual, judgments, such as how accurate it would be possible to be.

For additional information, see: Mobasseri, S., Stein, D. H., & Carney, D. R. (2022). The accurate judgment of social network characteristics in the lab and field using thin slices of the behavioral stream. *Organizational Behavior and Human Decision Processes*, *168*, Article 104103. https://doi.org/10.1016/j.obhdp.2021.09.002

References

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Bernard, H. R., Killworth, P. D., Sailer, L., & Kronenfeld, D. (1984). The problem of informant accuracy: The validity of retrospective data. *Annual Review of Anthropology*, *13*, 495–517.

Brewer, M. B. (1997). On the social origins of human nature. In C. McGarty & S. A. Haslam (Eds.), *The message of social psychology* (pp. 54–62). Oxford, UK: Blackwell.

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

Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review*, *1*, 276–298.

Dunbar, R. I. (2008). Cognitive constraints on the structure and dynamics of social networks. *Group Dynamics Theory and Research Practice*, *12*, 7–16.

Durkheim, É. (1995). *The elementary forms of religious life,* translated by Fields, K. E. New York: Free Press.

Ibarra, H., Kilduff, M., & Tsai, W. (2005). Zooming in and out: Connecting individuals and collectivities at the frontiers of organizational network research. *Organization Science*, *16*, 359–371.

Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, *61*, 587–610.

Lin, N. (2001). *Social capital: A theory of social structure and action.* New York, NY: Cambridge University Press.

Moreno, J. L. (1934). Who shall survive? A new approach to the problem of human interrelations*. Nervous and Mental Disease Publishing Co*, 58, 2–20.

Perry, B. L., Pescosolido, B. A., & Borgatti, S. P. (2018). *Egocentric network analysis*. Cambridge: Cambridge University Press.

Rossi, P. H. (1966). Research strategies in measuring peer group influence. In T. M. Newcomb & E. K. Wilson (Eds.), *College peer groups* (pp. 190–214). Chicago: Aldine.

Wellman, B. (1993). An egocentric network tale: Comment on Bien et al. 1991. *Social Networks*, *15*(4), 423–436. https://doi.org/10.1016/0378-8733(93)90015-D

Ecological Validity of Your Study: Introducing the Multidimensional Assessment of Research in Context (MARC) Tool

Sandra Naumann[1], Michelle L. Byrne[2,3], Alethia de la Fuente[4,5,6], Anita Harrewijn[7,8], Tehila Nugiel[9], Maya Rosen[10], Nienke van Atteveldt[11], & Pawel J. Matusz[12,13,14,15]

✉ pawel.matusz@gmail.com

[1] Berlin School of Mind and Brain, Department of Psychology, Humboldt-Universität zu Berlin, Germany

[2] Department of Psychology, University of Oregon, United States

[3] Turner Institute for Brain and Mental Health, School of Psychological Sciences, Monash University, Australia

[4] Buenos Aires Physics Institute (IFIBA) and Physics Department, University of Buenos Aires, Argentina

[5] Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Argentina

[6] National Scientific and Technical Research Council (CONICET), Argentina

[7] Emotion and Development Branch, National Institute of Mental Health, United States

[8] Department of Psychology, Education, and Child Studies, Erasmus University Rotterdam, The Netherlands

[9] Department of Psychology, The University of Texas at Austin, United States

[10] Department of Psychology, Harvard University, United States

[11] Section of Clinical Developmental Psychology, Research Institute LEARN!, Institute of Brain and Behavior, Faculty of Behavioral and Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands

[12] The MEDGIFT Lab, Information Systems Institute, School of Management, University of Applied Sciences & Arts Western Switzerland (HES-SO Valais), Sierre, Switzerland

[13] Institute of Health Sciences, School of Health Sciences, HES-SO Valais-Wallis, Sion, Switzerland

[14] Department of Hearing and Speech Sciences, Vanderbilt University, United States

[15] The Sense Innovation and Research Center, Lausanne and Sion, Switzerland

In this paper we present the Multidimensional Assessment of Research in Context (MARC) Tool. The MARC Tool is the result of our discussions about ecological validity and representative design of experiments. In cognitive neurosciences, fundamental principles of mental processes and functional brain organization can be studied with different approaches (or contexts), using highly controlled tasks to real-world research designs. These approaches have recently been integrated within a cyclical framework, composed of three "nodes": (1) controlled laboratory approach, (2) partially naturalistic laboratory approach, and (3) naturalistic real-world research approach (Matusz et al., 2019). All three nodes are of critical importance to creating more ecologically valid research. Studying a neurocognitive process of interest, e.g., directing attention to currently important events in the environment, with all three approaches is important because only together they can provide a more complete understanding of the process of interest. They work so well, because each approach can test a specific type of hypotheses and generate new ones for further investigation. For example, a traditional lab approach focuses on hypotheses teasing apart neural mechanisms responsible for processing discrete, isolate task aspects in different senses, while the more naturalistic approaches enable us to probe mechanisms orchestrating neurocognitive functions in more typical situations where does occur than in traditional laboratory context. Based on this theoretical framework, we have developed the MARC tool to help researchers more easily

delineate the approach that they have taken in their study (see below). Identifying the approach to which the study effectively belongs to is an important step towards improving comparisons between studies, which in turn should lead to more accurate theories of mental functions.

The MARC Tool provides a way to describe the degree of ecological validity of each component of a study. First, the researcher will describe the behavior they plan to observe and the context they aim to generalize to. Then, the tool poses seven questions concerning the study's characteristics: sample, testing site, task, stimuli, measures, non-research stakeholders, and a (potential) intervention part of the study. Each of these questions should be answered on a 3-point scale linked to each of the three approaches: controlled laboratory-based, partially naturalistic, and naturalistic real-world. The MARC Tool includes examples on each of these three levels. The output of the MARC Tool is a summary of the questions and a compass plot, reflecting the level(s) of the study's ecological validity. This output can be used for presentations, pre-registrations, grant proposals, and papers. With this tool we hope to improve drawing conclusions across studies and raise awareness about the importance of generalizability of research findings.

The MARC Tool is freely available online at: https://marcform-git.herokuapp.com

References

Matusz, P. J., Dikker, S., Huth, A. G., & Perrodin, C. (2019). Are we ready for real-world neuroscience? *Journal of Cognitive Neuroscience*, *31*(3), 327–338. https://doi.org/10.1162/jocn_ e_01276

Unwittingly Representative: Development of Stimulus Cases for a Medical Differential Diagnosis Tutorial

Frank J. Papa, DO, PhD

Texas College of Osteopathic Medicine, United States

✉ frank.papa@unthsc.edu

Triggered by the announcement for last year's Brunswik Society meeting, I found my copy of The Essential Brunswik and began to go through it. I felt relieved to read Hammond's introduction, wherein he describes Brunswik's work as largely being complex/dense and off putting because of his writing style and precision/obliqueness of thought. Nonetheless, as I slugged through a few chapters I came to think that my work, involving the use of an artificial intelligence research tool (KBIT) to model, simulate, and explore the roles of differential diagnosis (DDX) oriented knowledge-base structures and cognitive processes (inferencing mechanisms), also approximated what Brunswik called the incorporation of "Representative Design." That is, it addressed Brunswik's belief that a researcher's experimentation should also be designed to carefully represent the population of learning targets/objects comprising the task environment.

Translated into the world of differential diagnosis (DDX) research, the set of training and testing cases used for education should represent both: 1) a sampling of the common and/or important diseases (e.g., myocardial infarction, pulmonary embolus, dissecting thoracic aortic aneurysm, pneumonia, etc.) likely to cause a given clinical problem (e.g., acute chest pain), and 2) the "polymorphic nature" of each disease's "presenting" signs and symptoms (i.e., different combinations of signs and symptoms with which patients suffering from each disease might present to the clinician for a diagnostic evaluation).

At the heart of my research with KBIT was the assumption that an individual's knowledge of the frequency with which signs and symptoms are associated with a given disease (i.e., knowledge in the form of disease by feature frequency sensitivity/probability estimates) was a critical factor in diagnostic performance/accuracy. But also having a robust panel of clinically confirmed chest pain test cases was an important element of my KBIT investigations.

I had two means of constructing a sample population of chest pain cases for educational use: 1) collecting real cases and their associated clinical findings (I had 101 real chest pain cases representing nine common and/or important diseases likely to cause a patient to present for an evaluation of their chest pain), and 2) generating test cases from experts' judgments of how frequently each disease might manifest in a given patient in terms of their characteristic signs and symptoms. KBIT could generate a variety of such cases via the use of a Monte Carlo procedure (case generation guided via the experts' chest pain disease by feature conditional probability estimates and KBIT's case generation algorithm). While the Monte Carlo process may produce a set of cases with the appropriate prevalence of cues, the cue co-occurrences are not necessarily representative of the ecology. In contrast, a sample of actual cases would be representative in both ways.

However, the process did not produce a third form of representativeness. That is, did not assure that the variety of case portrayals chosen to represent a given disease in the learning set approximated the prevalence of diseases in a population of real cases presenting with the same disease. Instead, the number of cases used in my KBIT research was determined by the need to cover both common and important (life threatening) diseases (even if the important diseases might be rare), and by KBIT's training strategy: an instructional strategy designed to enable students to differentiate/discriminate close disease competitors (i.e., diseases that are most frequently confused with each other due to overlaps in the signs and symptoms that they diseases usually have).

Early studies provided me with reason to suspect that KBIT's successful simulation of an expert's superior diagnostic performance (compared to medical novices) was predicated upon knowledge of disease by feature frequency or sensitivity estimates; a knowledge base that likely represented a more valid portrayal of the range of case portrayals possible for a given disease (Papa & Meyer, 1987). In a subsequent study I found that students who experienced supplemental chest pain diagnosis training via numerous practice cases (via the KBIT tutorial) were able to construct disease by feature frequency matrices that correctly diagnosed more cases in KBIT's case bank, compared to the disease by feature frequency estimates collected from the same students at the beginning of their clinical rotations (Papa & Meyer, 1989). Thus, the construction of a training experience that was designed to mirror a large sample of chest pain cases would lead to changes in their disease by feature frequency estimates (as the result of a broader exposure to that external environmental reality ... exposure to a more representative set of case portrayals for each disease comprising a given problem space). Did this mean I had more direct evidence that improvements in students' diagnostic performance are in part related to the amount of exposure to the target environment?

Is what I was doing, early on, an example of what Brunswik meant by Representative Design? Although it was done to study judgment, per se, but rather to provide an educational experience, still it seems to show the benefit in using stimulus sets/training materials where the target environment is accurately represented.

References

Papa, F. J., & Meyer, S. (1987). An expert program shell designed for extracting "disease prototypes" and their use as models for exploring the "strong problem solving methods" employed in clinical reasoning. In I. R. Hart (Eds.), *Further developments in assessing clinical competence* (354–364). Quebec, Canada, Heal Publications.

Papa, F. J., & Meyer, S. (1989). A computer-assisted learning tool designed to improve problem solving skills. *Annals of Emergency Medicine*, *18*(3), 273–296. https://doi.org/10.1016/s0196-0644(89)80412-3

Generalization Bias: An Overlooked Cause of Overgeneralizations in Psychology

Uwe Peters

Leverhulme Centre for the Future of Intelligence, University of Cambridge, United Kingdom

Center for Science and Thought, University of Bonn, Germany

✉ up228@cam.ac.uk

Reliable experimental studies produce findings that can be extrapolated from the particular study context to other contexts. Yet, many researchers have found that a large proportion of psychological studies lack this feature, as they are conducted only in laboratory settings (Shamay-Tsoory & Mendelsohn, 2019), involve predominantly only students (Hanel & Vione, 2016), or sample mostly only WEIRD (western, educated, industrialized, rich, democratic) populations (Rad et al., 2018). This means that the studies' representativeness and so generalizability is low.

Brunswik (1947), too, highlighted this problem and influentially distinguished between "populational" generality of results, referring to the applicability of results across new participants, and "situational" (or "ecological") generality, referring to the applicability of results across related but new stimuli or study material (e.g., photos of persons as opposed to films or face-to-face encounters). Brunswik noted that especially situational generality was missing in much experimental research, and researchers failed to qualify their results accordingly, resulting in "numerous overgeneralizations" (1947, p. 192). The problem that many psychologists tend to ignore variation in stimulus sampling when drawing their conclusions, aka the "stimulus-as-fixed-effect fallacy" (Clark, 1973), is thought to drastically inflate false-positive rates, particularly with large participant samples and small stimulus samples, as stimulus variance is then high (Yarkoni, 2021).

While Brunswik and others have done much to highlight many studies' inadequate populational and situational representativeness and the problem of scientific overgeneralizations, they have left the *causes* of them largely unexplored. In a recent paper, my colleagues and I (Peters et al., 2022) aim to fill this research gap by looking at the psychology of psychologists themselves and situating it within the institutional and social structures of science in which psychologists operate.

It is a common, typically tacit assumption that when scientists generalize their results, this is a conscious, fully controlled ("system 2") process. Integrating findings from across the cognitive sciences, we argue against this notion and propose that psychologists' extrapolations often involve a "generalization bias." This is a cognitive ("system 1") tendency that operates (in the absence of interventions) automatically and frequently leads people to unintentionally generalize from some experiences, including study results from particular samples, to all experience (of that kind) or broader populations even when the evidence does not warrant it. The outcome, in psychological studies, is (*inter alia*) precisely the overgeneralizations and neglect of populational and situational generality of results that Brunswik highlighted.

Our view takes seriously recent contributions arguing that to fix methodological problems in science (e.g., the "replication crisis"), social factors such as scientific training and incentives are

important but "we need also to take into account […] human cognitive biases", as "scientific thinking is not natural for humans: To be good scientists, we often have to actively inhibit our normal ways of thinking" (Bishop, 2019).

Our argument involves the following main steps. We first review studies that found that participants more readily encoded more general propositions about whole categories of individuals ("people", "women", etc.) than quantified propositions (e.g., "*some* people", "*many* women", etc.) and systematically misremembered quantified statements as generalized ones independently of their cognitive resources and against their goal to be accurate. Participants also needed only very little evidence about some features of some members of an unfamiliar group to form the generalization that the group as a whole had these features. Yet, when presented with the same generalization first, they took it to be supported by much more evidence. These findings suggest that participants had an unwittingly operating automatic generalization tendency, i.e., a bias.

In a next step, we relate these data to scientists, i.e., a population not tested in the original studies. Since empirical scientists, including psychologists, are especially trained to tailor their claims to their evidence, they may be immune to a generalization bias. However, we review evidence of, for instance, pervasive overgeneralizations from Western, educated, industrialized, rich, democratic (WEIRD) population samples in psychology articles to argue that psychologists, too, are likely vulnerable to a generalization bias in their scientific induction.

These overgeneralizations may be explained by other factors than a cognitive bias such as psychologists' forgetfulness of their training, methodological ignorance, or social influences: Broader claims may be easier to publish, as they suggest a higher importance of the results, and publication pressure on researchers may thus contribute to overgeneralizations. However, scientific overgeneralizations can even be found in articles whose authors are fully aware of the problem they pose and know how to tackle them. For instance, Henrich et al. (2010, p. 1) begin their influential critique of WEIRD population sampling and many psychologists' overgeneralizations with the overgeneralization: "Behavioral scientists routinely publish broad claims […]". This suggests that it isn't forgetfulness or ignorance that is driving the problem but an unintentionally operating, automatic generalization bias. This view also aligns well with the earlier mentioned empirical evidence of such a bias in laypeople's thinking.

And importantly, for instance, publication pressure, journal guidelines requiring condensed formats (promoting qualifier omission), and so on are not competing explanations to the generalization bias account of scientific overgeneralizations. Rather, they in fact become more plausible if it is granted that this bias, too, can also affect psychologists. This is because including this bias as a possible cause allows granting that researchers or editors may not deliberately but inadvertently contravene epistemic standards when producing or encouraging overgeneralizations. Conversely, the notion that, for instance, publication pressure may promote overgeneralizations in turn can make it more plausible that psychologists, too, may have, or may increasingly develop, a generalization bias, as such pressure can entrench a habitual generalization tendency. We thus argue that external or structural, i.e., social and institutional, as well as psychological factors interact and together fuel scientific overgeneralizations.

This proposal is novel, because even though other researchers have highlighted several structural causes of generalizability issues in psychology, it has not been considered yet that a cognitive

generalization bias may play a significant role also. The dynamic interplay between this internal factor and the external structural causes of overgeneralizations has gone unnoticed also. But recognizing that a generalization bias may be at play and that an interaction between internal and external factors may promote the generalizability problems in psychology highlighted by Brunswik and others is likely important for fixing these problems. This is because mitigation strategies will then need to target both factors to be fully effective.

While changing the broader structures of science that facilitate overgeneralizations (publication pressure, competition, etc.) may be difficult, other concrete counterstrategies exist. We recommend (*inter alia*) that researchers consider adopting a personal checklist to mitigate generalization bias. This list may include reminders to include relevant qualifiers (e.g., "some", "many") in conclusions, or consider making "minimum statements of the form 'at least in certain individual cases […]'", as Brunswik (1947, p. 194) suggests. We recommend that this list also include reminders to use the past tense when describing findings (which automatically restricts claims to specific groups), and to ensure that both the participant sample and stimuli sample are representative.

More generally, merely changing the way we, as social scientists, think about the nature of scientific induction so that we no longer view it as a fully conscious and voluntary but as a partly automatic, generalization-bias-driven process can already help reduce overgeneralizations. After all, it may make us more alert to examining the scope of the conclusions we draw, which can then prompt us to add important qualifications.

References

Bishop, D. (2019). Fixing the replication crisis: The need to understand human psychology. *Psychological Science*. URL: https://www.psychologicalscience.org/observer/fixing-the-replication-crisis-the-need-to-understand-human-psychology

Brunswik, E. (1947). *Systematic and representative design of psychological experiments: With results in physical and social perception*. Berkeley, CA: University of California Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*(4), 335–359.

Hanel, P. H. P., & Vione, K. C. (2016). Do student samples provide an accurate estimate of the general public? *PLoS ONE*, *11*(12), e0168354. https://doi.org/10.1371/journal.pone.0168354

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83.

Peters, U., Krauss, A., & Braganza, O. (2022). Generalization bias in science. *Cognitive Science*, *46*(9), e13188. https://doi.org/10.1111/cogs.13188

Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of Homo sapiens: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 11401–11405.

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.

Yarkoni, T. (2021). The generalizability crisis. *Behavioral and Brain Sciences*, *45*, E1. https://doi.org/10.1017/S0140525x20001685

Insurance Decisions of Doctors in Private Practice in Germany: Dissertation Research Status and Approach

Claudia Pitterle

Doctoral School of Economics and Business Administration, University Alexandru Ioan Cuza din Iasi, Romania

✉ pitterle.c@gmail.com

The insurance demand of private households has occupied researchers since the late 1960s. At that time, they were concerned with the decision-making process in selecting an optimal deductible in comprehensive motor vehicle insurance (Murray, 1971). The deductible is the premium that a policyholder is willing to pay in the event of a claim.

It quickly became apparent that policyholders do not behave rationally when asking for insurance coverage. In other words, how one should decide according to expected utility theory as a rational decision maker or homo economicus. Emotions or information perception errors do not occur in a utility maximizer. Focus of economic sciences became more and more on how behavior actually showed up.

Human decisions can be influenced by different cognitive biases. To cope with the flood of information perception and the limited information perception capacities, humans use mental shortcuts or rules of thumb, so-called heuristics. Decisions are also influenced by the framework of information presentation and by one's emotions, and not only according to one's own benefit (Richter et al., 2017).

Countless summaries of research and books from behavioral economics, which would go beyond the enumeration scope here, have been published since then. Some interesting and relevant publications can be found at: Jaspersen (2015); Kunreuther et al. (2013); Kunreuther and Pauly (2004); Richter et al. (2017); Theil (2002).

Such non-rational behavior, as described above, is also evident in my day-to-day professional consulting work as a Certified Financial Planner[1], CFP® when advising on insurance products. Insurances, which secure the private as well as professional existence and protect against economic ruin, are not contracted by customers or are pushed into the future and then simply forgotten. Also, the entrepreneurial power of attorney is of great importance for an entrepreneur or in my field of research, a doctor in a single practice, in order to be able to appoint a substitute in case of longer absence due to illness or even death. The practice can thus continue without restrictions and practice revenues can continue to be generated.

In the fall of 2020, I started to explore this behavior in more detail as part of my dissertation and to work out explanatory attempts from the field of behavioral economics. In doing so, I am examining physicians in private practice in Germany who run their own practice or work in group practices. The only compulsory insurance for doctors in private practices in Germany is

[1] https://www.cfp.net/

professional liability insurance. This protects against the financial consequences of claims for damages and covers this existential risk. All other insurances, such as inventory insurance, business interruption insurance, cyber risk insurance, etc. are voluntary and the doctor decides himself/herself whether to take out an insurance policy.

Cyber-crime and attacks on the German healthcare system in hospitals and specifically pharmacies and medical surgeries have increased in recent years, see: Industry Report Cyber Risks at Doctors and Pharmacies (GDV, 2019). Patient data has become a form of prey. Damages related to hacker attacks, such as fraud or other acts of cybercrime are secured by cyber insurance. Hence, interesting questions and results from my research present themselves.

In the fall of 2020, I conducted in-depth interviews with five doctors in private practice in Munich and the surrounding area, who have been in private practice from one year to more than 20 years (Pitterle, 2021). Among other things, I clarified the following questions: Which insurances were contracted, how important are which insurances for the doctor and were they contracted? Are insurances updated and were there any claims? How is cyber risk insured, a topic currently represented in the media? Are there entrepreneurial powers of attorney?

The aim of the in-depth interviews was to find out whether doctors even knew which contracts they had contracted. Also, according to which criteria was insurance bought, such as basic protection, price-performance or to be "all-around carefree."

In summary, it turned out that doctors know little about their own insurance portfolio. Already with the inquiry over the insurance existence it came again and again to questions of understanding the contained insurance cover of a contract. There is also no congruence between the contracts perceived as important and those contracted. The issue of cyber risk was not recognized at all for one's own practice and the doctor feels protected, whereas the general risk in Germany of cyber threat was assessed as high. Powers of representation were not regulated for individual practices.

Due to the small number of interviews conducted, the results are not representative, but give an initial indication of decision-making behavior and thus the possibility of expanding the questions in an online survey.

As the dissertation progressed, I then developed a questionnaire that was rolled out widely via the Arzt & Wirtschaft[2] newsletter, "What's up doc?" podcast[3], and social media such as LinkedIn and Xing. Participation among doctors is very low, as generally suspected. Nevertheless, 80 complete surveys came in after about two months. These are currently being evaluated.

A very interesting approach from the perspective of the Brunswikian Lens (Brunswik, 1952) offers the assessment of the risk "Cyber Risk" on medical surgeries (GDV, 2019). What judgment does a doctor form through his lens regarding the cyber threat in Germany and his own practice?

[2] https://www.arzt-wirtschaft.de/finanzen/versicherungen/versicherungsschutz-fuer-ihre-praxis-das-sollten-sie-wissen/
[3] https://www.arzt-wirtschaft.de/doctolib-podcasts/schutz-fuer-arztpraxen-welche-versicherungen-lohnen-sich-wirklich/

Here, the Forsa survey from 2019 provides very interesting results, which can be explained by the weighting of cues, leading to a subjective judgment of the doctor. Forsa (Politik-und Sozialforschung GmbH) conducted a representative survey on the topic of "Cyber Risks in Healthcare" on behalf of the German Insurance Association (GDV).

In 200 medical surgeries and 101 pharmacies, the employees entrusted with the IT infrastructure or responsible for cybersecurity, were surveyed from June 11 to July 6, 2018, using computer-assisted telephone interviews.

At the beginning, the commissioners were asked how high they considered the risk for medical surgeries or pharmacies in Germany of becoming victims of cybercrime. More than half of the doctors, namely 53%, saw the risk as rather low/very low. When asked about their own practice's assessment of becoming a victim of cybercrime in the next two years, 81% of doctors saw the risk as low/very low. Reasons for judgment, which can be regarded as the Brunswikian cues, are shown Table 1.

Table 1. Percentage of reasons for believing a low risk of cybercrime*.

Belief that the risk for their own practice/pharmacy is very or rather low, because…	Practices (%)**	Pharmacies (%)**
…their computer systems are protected against cyber-attacks	80	89
…their practice/pharmacy has never previously been a victim of a cyber-attack	81	72
…patient/customer data is exclusively stored and sent encrypted	64	49
…their practice/pharmacy is too small to be of interest to cybercriminals	56	49
…they do not store patient/customer data on computers that are connected to the network/on the Internet	49	32
…their data is not of interest to cybercriminals	45	37
…patient/customer data is not locally stored, only in the cloud	22	8

***** Basis: Respondents who see a rather or very low risk for their own practice/pharmacy to become victims of cybercrime in the next two years.

****** Percentage sum greater than 100, since multiple answers may be possible.

It is interesting to see how the weighting of the cues turned out in relation to the attitude of the risk to one's own practice. These cues, weighted by the doctors, led to the judgment that their own practice is less threatened by cybercrime than the general threat situation in Germany. Thus, the weighting of cues leading to the judgment of the overall cyber risk threat situation in Germany must be different than that on one's own practice.

The information retrieved in the brain and the compilation into a judgment, i.e., the judgment policy, leads to a different perception of risk. And the compilation and weighting of cues also leads

to the fact that cyber insurance, which protects in the event of an attack, is not important and has not been contracted. From my perspective as a CFP® and that of my colleagues, cyber insurance was rarely if ever requested by doctors in 2018.

Unfortunately, asking German-based insurance companies in the spring of 2022 how much cyber insurance has been contracted in recent years is proving very difficult, as this data may not be released. However, an increase in demand and an increase in the number of policies taken out have become apparent in recent years, and this could be communicated. Since this insurance segment is also still very young, the GDV does not have a separate segment for counting it in Germany.

I have integrated this survey into my questionnaire because I am interested in whether the individual cyber threat risk has changed over the years and whether there have been more cyber insurance policies taken out. The data is currently being analyzed. It will therefore be exciting to see how the cues are weighted and combined to form a verdict.

References

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

GDV. (2019, May 31). *Branchenreport Cyberrisiken bei Ärzten und Apotheken*. www.gdv.de. Retrieved 21 August 2020, from https://www.gdv.de/resource/blob/45196/ae262d6702e2d9f5446c780a22450d23/download-branchenreport-cyber-aerzte-und-apotheker-data.pdf#page=1&zoom=70,-466,853

Jaspersen, J. G. (2015). Hypothetical surveys and experimental studies of insurance demand: A review. *Journal of Risk and Insurance*, *83*(1), 217–255. https://doi.org/10.1111/jori.12100

Murray, M. L. (1971, September). A deductible selection model. Development and application. *The Journal of Risk and Insurance*, *38*(3), 423. https://doi.org/10.2307/251404

Kunreuther, H., & Pauly, M. (2004). Neglecting disaster: Why don't people insure against large losses? *Journal of Risk and Uncertainty*, *28*(1), 5–21. https://doi.org/10.1023/b:risk.0000009433.25126.87

Kunreuther, H. C., Pauly, M., & McMorrow, S. (2013). *Insurance and Behavioral Economics: Improving Decisions in the Most Misunderstood Industry* (Illustrated). Cambridge University Press.

Pitterle, C. (2021, March 30). How do doctors evaluate insurance for their surgery and what decisions are made for or against taking out insurance? *International Journal of Management and Humanities*, *5*(7), 9–15. https://doi.org/10.35940/ijmh.g1249.035721

Richter, A., Ruß, J., & Schelling, S. (2017). *Moderne Verhaltensökonomie in der Versicherungswirtschaft: Denkanstöße für ein besseres Verständnis der Kunden (essentials) (German Edition)* (1. Aufl. 2018). https://doi.org/10.1007/978-3-658-19841-1

Theil, M. (2002). *Versicherungsentscheidungen und Prospect Theory*. Springer Publishing.

On the Ecological Validity of Soundscape Reproduction in Laboratory Settings

Cynthia Tarlao & Catherine Guastavino

School of Information Studies, McGill University and Centre for Interdisciplinary Research in Music Media and Technology, Canada

✉ cynthia.tarlao@mail.mcgill.ca

The concept of ecological validity was first introduced by Egon Brunswik (1943) and further developed by Gibson (1979) in the context of visual perception. This ecological approach was later applied to auditory perception with the study of everyday sounds and soundscapes, defined as the collection of sounds experienced in an environment.

The ever-growing body of soundscape research includes studies conducted in everyday life environments and in laboratory settings (e.g., Misdariis et al., 2019; Skoda et al., 2019). Yet laboratory settings differ from everyday life situations and therefore may elicit different perceptions. The present study explores the *ecological validity* (or in current readings of Brunswikian theory, *representativeness*) of soundscape reproduction in laboratory settings using Ambisonics (a spatial audio recording and reproduction technique [Gerzon, 1973]) and comparing two different modes of questionnaire administration. Soundscape evaluations were measured using scales derived from the literature. These include the Swedish Soundscape Quality Protocol (SSQP), developed in Swedish and English in lab-based experiments (Axelsson et al., 2010), to collect soundscape evaluations along two main dimensions of *pleasantness* and *eventfulness*. We also measured soundscape *appropriateness*, which is understood as soundscape *appropriateness* for specific activities (Steele et al., 2019) in this study. In addition, we assess the potential for restoration provided by urban soundscapes, using the recently developed Perceived Restorativeness Soundscape Scale (Payne & Guastavino, 2018). Our study focused on the soundscapes of a public square in downtown Montreal (Fleurs-de-Macadam square).

The ecological validity of an experimental design rests on three elements: 1) the participants being representative of the population the results are intended to be generalized to; 2) the experimental conditions being representative of the actual conditions the results are meant to apply to; and 3) the task (including instructions and data collection instruments) eliciting similar processes than in the everyday life situations (Brunswik, 1943).

Regarding the first requirement related to participants, residents who use the public square on a regular basis were recruited through the neighborhood's newsletter, ensuring that participants were representative of the population of interest and familiar with the soundscapes reproduced. Indeed previous work has shown that sound experts (sound engineers) and non-expert city users pay attention to different aspects of soundscape reproduction, highlighting the relationship between individual experience and ecological validity (Guastavino, 2003, 2009).

Regarding condition representativeness, different reproduction methods and systems can be preferred depending on the soundscape reproduced (Guastavino et al., 2005; Guastavino & Katz,

2004). A recent study found no significant differences between *in situ* and Ambisonics reproduction in terms of SSQP ratings and dominant sound sources (Hong et al., 2019). Finally, in terms of the experimental settings, different reproduction systems prompt different cognitive representations and Ambisonics was found to elicit strong immersion while allowing for sound source identification (Guastavino et al., 2005). Hong et al. (2019) also found that Ambisonics provided higher immersion, realism, and externalization than binaural reproduction methods. Another aspect of the experimental process is the procedure – here specifically, the mode of questionnaire administration. In general, soundscape questionnaires are administered *in situ* with pen and paper, while laboratory studies are more conducive to computer-based tasks. It is therefore fundamental to explore the transferability of results from one mode of administration to the other in the context of soundscape studies.

These last two requirements are the focus of this paper: 1) validating the conceptualizations underlying soundscape evaluations and 2) investigating the influence of the mode of administration of the questionnaire and of the excerpt of recording reproduced on judgments.

To answer the research questions, this study was structured in two connected parts. First, data was collected in the public space through a) users' questionnaire-based soundscape evaluations (N = 185, 102 women, 76 men, age = 34.8 ± 14.8) and b) audio recordings taken during a representative portion of the data collection periods. Second, samples were selected from the audio recordings to be representative of the different conditions[1] described below, based on the summer-long experience of the authors on site, including prototypical sounds, e.g., of space users, passersby, traffic, businesses nearby, nature, and excluding salient sound events such as firetruck sirens and construction. The selected samples were reproduced in a laboratory setting on a 3-dimensional array of 17 loudspeakers using Ambisonics to collect participants' soundscape evaluations (N = 34, 18 women, 16 men, age = 45.1 ± 16.8). Conditions were selected in a factorial design, with 2 locations (quiet side facing residences vs. noisy side along a busy street) × 2 days of the week (weekday vs. weekend) × 2 times of day (afternoon vs. evening) × 2 excerpts (selected 2-minute samples within each recording), for a total of 16 stimuli. We then compared the evaluation collected in the laboratory and on-site.

First, our findings confirm that participants hold similar conceptualizations of the underlying soundscape dimensions, as demonstrated by a confirmatory factor analysis (CFA), both *in situ* and in the laboratory. Second, based on the validated CFA results, we found scalar measurement invariance between the two modes of administration tested in the laboratory (computer-based and pen-and-paper questionnaires). In other words, participants understand and use the soundscape scale in a similar way with computer-based and pen-and-paper modes of administration. Additionally, no differences were observed between different excerpts selected from a 10-minute recording of the same condition. Finally, semi-parametric MANOVA and follow-up ANOVA reveal similar patterns between the *in situ* and the laboratory results: the quiet side is judged more pleasant and less eventful than the noisy side. These results, in line with previous studies research (Davies et al., 2014; Guastavino et al., 2005, 2007), point to the ecological validity of Ambisonic

[1] Note that this selection cannot guarantee that the distribution of samples is representative of on-site situations (e.g., space users might spend more time on the quiet side of the space).

reproduction for soundscapes, when comparing each stimulus to the corresponding condition on-site.

An interesting finding that we did not foresee is that on-site evaluations seem to be much more positive than laboratory evaluations. This could be attributed to several, potentially overlapping reasons. First, sensory experiences on site integrate other sensory modalities which could alleviate the unpleasantness of city noise. Second, place attachment within the neighborhood (e.g., historical significance, sense of community) may increase user satisfaction and in turn, improve soundscape pleasantness. Third, city users might expect the city to be noisy and therefore employ conscious coping strategies to mitigate said noise – such as using the space at specific times. Finally, the immersive reproduction of traffic noise could be an uncomfortable reminder of how pervasive traffic is in the city by making it harder to ignore in a laboratory setting.

These results point to the need to understand the exact ways in which Ambisonic soundscape reproduction differs from on-site experiences. However, the present results reveal highly similar soundscape conceptualizations in laboratory setting and on site despite differences in variability across participants, justifying the adoption of Ambisonics reproduction for soundscape research.

For further details, please read the full paper at: Tarlao, C., Steele, D., & Guastavino, C. (2022). Assessing the ecological validity of soundscape reproduction in different laboratory settings. *PLOS ONE*, *17*(6), e0270401. https://doi.org/10.1371/journal.pone.0270401

References

Axelsson, Ö., Nilsson, M. E., & Berglund, B. (2010). A principal components model of soundscape perception. *The Journal of the Acoustical Society of America*, *128*(5), 2836–2846. https://doi.org/10.1121/1.3493436

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*, *50*(3), 255–272. https://doi.org/10.1037/h0060889.

Davies, W. J., Bruce, N. S., & Murphy, J. E. (2014). Soundscape reproduction and synthesis. *Acta Acustica United with Acustica*, *100*(2), 285–292. https://doi.org/10.3813/AAA.918708.

Gerzon, M. A. (1973). Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society*, *21*(1), 2–10.

Gibson, J. J. (1979). *The Ecological approach to visual perception*. Boston, Massachusetts: Houghton, Mifflin and Company.

Guastavino, C. (2003). *Étude sémantique et acoustique de la perception des basses fréquences dans l'environnement sonore urbain*. Ph.D. Paris 6, Paris.

Guastavino, C. (2009). Validité écologique des dispositifs expérimentaux. In: Dubois D (ed.) *Le Sentir et le Dire*. Paris: L'Harmattan, pp. 233–252.

Guastavino, C., & Katz, B. F. G. (2004). Perceptual evaluation of multi-dimensional spatial audio reproduction. *The Journal of the Acoustical Society of America*, *116*(2), 1105–1115. https://doi.org/10.1121/1.1763973.

Guastavino, C., Katz, B. F., Polack, J. D., Levitin, D. J., & Dubois, D. (2005). Ecological validity of soundscape reproduction. *Acta Acustica united with Acustica*, *91*(2), 333–341.

Guastavino, C., Larcher, V., Catusseau, G., & Boussard, P. (2007). Spatial audio quality evaluation: Comparing transaural, ambisonics and stereo. In *Proceedings of the 13th International Conference on Auditory Display*, Montreal, Canada, June 2007, pp. 53–59. Georgia Institute of Technology. Available at: https://smartech.gatech.edu/handle/1853/50032 (accessed 9 September 2019).

Hong, J. Y., Lam, B., Ong, Z. T., Ooi, K., Gan, W. S., Kang, J., ... & Tan, S. T. (2019). Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications. *Building and Environment*, *149*, 1–14. https://doi.org/10.1016/j.buildenv.2018.12.004.

Misdariis, N., Marignier, L., Dianoux, C., et al. (2019). Experimental validation of a urban traffic noise annoyance model. In *Proceedings of the 26th International Congress on Sound and Vibration*, Montreal, Canada, p. 8.

Payne, S. R., & Guastavino, C. (2018). Exploring the validity of the perceived restorativeness soundscape scale: A psycholinguistic approach. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02224.

Skoda, S., Becker-Schweitzer, J., & Kiene, J. F. (2019). Evaluation of urban sound environments in-situ and in laboratory. In *Proceedings of the 26th International Congress on Sound and Vibration.* Montreal, Canada, p. 3.

Steele, D., Bild, E., Tarlao, C., & Guastavino, C. (2019). Soundtracking the public space: Outcomes of the musikiosk soundscape intervention. *International Journal of Environmental Research and Public Health*, *16*(10), 1865. https://doi.org/10.3390/ijerph16101865.

Using Clinical Judgement Analysis to Uncover the Implicit Clinical Decision-Making Policies of Pharmacists: A Multiple-Study Project

Janique Waghorn

Department of Pharmacy, Institute of Pharmaceutical Science, King's College London, United Kingdom

✉ janique.smit@kcl.ac.uk

Tim Rakow

Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, United Kingdom

Jennifer M. Stevenson

Department of Pharmacy, Institute of Pharmaceutical Science, King's College London, United Kingdom

Background

Pharmacy is a profession increasingly on the frontline of patient care, which is stepping out of the backroom. For example, the UK National Health Service (NHS) Long Term Plan published in 2019 states that pharmacists have an important role in improving population health-related outcomes. With rising numbers of pharmacists becoming prescribers, the expanded role of pharmacists in General Practice as well as a more clinical shift in the role of the community pharmacist, pharmacists initiating treatment plans is becoming part of common practice in the UK. The more traditional role of correcting prescribers and making recommendations is shifting to the role of decision-maker. It is therefore necessary to determine whether the initial education and training of pharmacists is still appropriate. A fuller picture of the decision-making processes of pharmacists is therefore required. To this end, I am using Clinical Judgment Analysis (CJA) to understand pharmacy decisions, based on Hammond's development of Brunswik's work (Hammond, 1996).

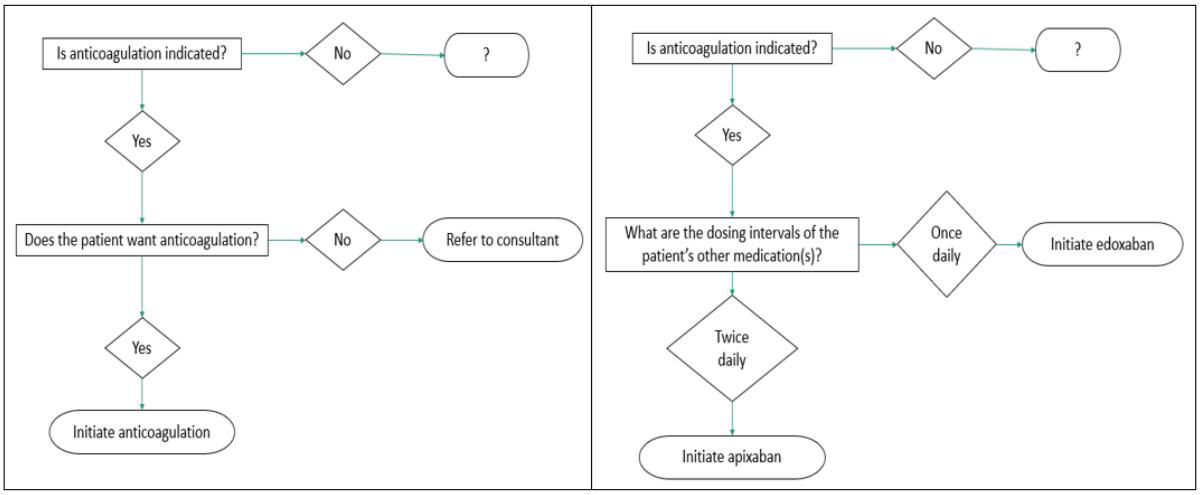
As part of a doctoral project, the overall aim is to demonstrate how CJA may be adapted to understand pharmacy decisions. The first part of this project used an illustrative study to outline how to apply CJA to pharmacists' decision-making and is outlined below. The decision to initiate anticoagulation, alongside appropriate risk judgments, provides the context for my multiple-study investigation.

Initial Study

Expert anticoagulation pharmacists were interviewed to identify and define the cue variables involved in this decision. Decision tasks with sixty scenarios were developed to explore the effect of these cues on pharmacists' decision-making processes and distributed to participants for completion. For each scenario, each participating pharmacist made three judgments (stroke risk, bleed risk, patient's ability to manage their condition) and one drug prescription decision.

CJA using multiple linear regression produced models of good fit for the prediction of stroke risk judgment and bleed risk judgment made by individual pharmacists. The cues chosen did not appear to systematically influence the judgment of how well the patient would manage their condition. Comparing participants' individual CJA models for their judgment of risk (both stroke and bleed) revealed differences between pharmacists in cue use and in how much their judgment agreed with the risk predicted by the literature.

The prescription decision was often captured well by a stepwise (lexicographic) model where the presence of one factor dictates the decision-making, rather than being based on a model where many variables are considered (see Figure 1). This is consistent with the way pharmacists are trained. Standard operating procedures are a legal requirement in UK pharmacies, promoting a stepwise approach for technical pharmacy tasks (e.g., dispensing). Possibly this training for the technical aspects of pharmacy has influenced how clinical decisions are made (including those not governed by standard operating procedures that mandate a stepwise approach).



Participant's 01 decision making policy

Participants' 02 and 04 decision making policy

Figure 1. Example decision models from a judgment analysis study reported in Waghorn et al. (2021).

The data from this study demonstrates that CJA generates insights into the clinical decision-making processes of pharmacists not uncovered by the current literature. This provides a springboard for more in-depth explorations; explorations that are vital to the understanding and ongoing development of the role of pharmacists. A fuller report the initial data for this study has been published (Waghorn et al., 2021). That paper also includes discussion of how and why CJA is a valuable tool for understanding the clinical decision-making of pharmacists. Further studies in this multiple-study project are ongoing and in preparation. These include analyses of in-clinic anticoagulation prescription decisions by pharmacy teams, and the development of video scenarios than can be used for pharmacist training and research using CJA.

Next Steps

The rest of this doctoral project will attempt to complete the following studies, based on this initial work:

- Study 01: CJA using designed scenarios in the context of oral anticoagulation initiation in atrial fibrillation (MPharm Students and PG IP students at a London University)
- Study 02: Applying CJA to real-life patient cases, a retrospective study of the initiation of DOACs in atrial fibrillation patients in a pharmacist-led anticoagulation clinic at a large London teaching hospital
- Study 03: Applying CJA to real-life patient cases, a retrospective study of the decision to initiate DOACs in patients at a large London teaching hospital
- Study 04: Design and implementation of video scenarios, in the context of anticoagulation prescribing, to use in CJA for training purposes

References

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

Waghorn, J., Bates, I., Davies, J. G., Jubraj, B., Rakow, T., & Stevenson, J. M. (2021). Clinical judgement analysis: An innovative approach to explore the individual decision-making processes of pharmacists. *Research in Social & Administrative Pharmacy: RSAP*, *17*(12), 2097–2107. https://doi.org/10.1016/j.sapharm.2021.05.006

Personality Judgements Based on People's Laptop Stickers

Gregory D. Webster

Department of Psychology, University of Florida, United States

✉ gdwebs@ufl.edu

Jessica T. Campbell

Kinsey Institute, Indiana University Bloomington, United States

Brunswikian lens models (Brunswik, 1956) have long been applied to examine the cues people use to gauge others' personality traits (see also Nestler & Back, 2013). Broadly, these studies range from judging personality based on people's bedroom and offices (Gosling et al., 2002) to World of Warcraft players based on their avatars and usernames (Harari et al., 2015).

In our recent work (Campbell et al., 2022), we extended research on personality judgements to people's laptop stickers. In the U.S., university students frequently adorn their laptops with stickers. These stickers serve multiple functions, including cues to identities, interests, beliefs, social groups, and political stances. When their laptops are open, these stickers serve as billboards that advertise their owner's identities and interests to others nearby and may facilitate conversations based on common interests. We sought to examine the extent to which university students show both consensus and accuracy in their ratings of other students' Big Five personality traits based solely on viewing high-quality digital photos of their laptop covers, which included three or more stickers.

In such personality judgment studies, *consensus* reflects the extent of agreement across perceivers, whereas *accuracy* describes the association between perceiver judgments of target personality traits and a criterion measure of personality, such as target self-reports of personality. We did not set out to examine cues from individual stickers because we were interest in people's overall gestalt judgements of others' personality traits based on all the stickers present on their laptop covers. Nevertheless, on a post hoc basis, we examined two likely cues based on observable features—"Mac-ness" and number of laptop stickers. To assess Mac-ness, we had two research assistants (RAs) independently rate whether each laptop was a Mac, PC, or unknown (obscured logo); scores ranged from 0 (definitely a PC) to 1 (definitely a Mac) with increments of 0.25 reflecting levels of uncertainly. Number of laptop stickers was a simple count variable that we normalized using a natural log transformation.

Our target stimuli included digital photographs of laptops from 147 university students with three or more laptop stickers. Our perceivers were eight research assistants (RAs) who made independent, multi-item ratings of the 147 laptop photos with the objective of accurately judging the laptop owner's personality traits (1,139 laptop ratings [97%] out of a possible 1,176). The laptop owners rated only themselves, whereas the perceivers (eight of our lab's RAs) completed the same personality measures for the 147 laptops: (a) the 60-item Big Five Inventory 2 (BFI-2; Soto & John, 2017) and the four narcissism items from the Dark Triad Dirty Dozen (Jonason & Webster, 2010). In addition to assessing the Big Five personality traits (i.e., extraversion,

agreeableness, conscientiousness, negative emotionality, and open-mindedness), the BFI-2 also assesses three 4-item facets per trait. For example, open-mindedness is comprised of the facets aesthetic sensitivity, intellectual curiosity, and creative imagination. Because this imposed a burden on each of our eight RAs—who we asked to make 9,408 total ratings (64 items × 147 laptops)—we gave them two weeks to complete the task.

We compared two analytic methods—a traditional one in which ratings are collapsed across raters, and a new one in which variance across raters is accounted for as well as measurement error (i.e., cross-classified structural equation modeling [CC-SEM]; Claus et al., 2020; Nestler & Back, 2017). The traditional approach assesses accuracy by correlating targets' self-reported personality with the average of perceivers' personality scores for them. Although more analytically complex, the CC-SEM approach should yield more optimal results because it appropriately models both targets (laptops and their owners) and perceivers (raters) as random effects while accounting for measurement error using a latent-variable approach.

Results indicated adequate consensus—agreement among perceivers across their ratings of laptops, a type of interclass correlation coefficient (ICC)—for most traits and their facets, with the exception of negative emotionality and its facets, which had comparatively low consensus effects. In addition, traditional methods slightly overestimated consensus effects for Big Five traits (M = .32, SD = .14) and their facets (M = .26, SD = .11) relative to CC-SEM-based estimates for traits (M = .25, SD = .14) and their facets (M = .23, SD = .12).

Regarding accuracy—the extent to which perceivers' ratings of laptop owners' personality traits actually covaried with the self-reported personality trats of those laptop owners—CC-SEM results showed significant accuracy effects for (a) extraversion (only after controlling for Mac-ness and number of stickers) and its facet of assertiveness, (b) negative emotionality's emotional volatility facet, and especially (c) open-mindedness and two of its three facets: aesthetic sensitivity and intellectual curiosity. In addition, traditional methods slightly underestimated accuracy effects for Big Five traits (M = .18, SD = .07) and their facets (M = .14, SD = .08) relative to CC-SEM-based estimates for traits (M = .21, SD = .08) and their facets (M = .18, SD = .12).

We applied a modified lens model to assess the extent to which observers used Mac-ness and (log) number of laptop stickers as cues. This analysis took the form of a CC-SEM statistical (vs. causal) mediation model, where the direct link between laptop owners' self-reports and perceivers' ratings were "mediated" by cues—Mac-ness and number of stickers. We found evidence of cue effects only for one set of variables: Number of stickers served as a reliable cue for ratings of aesthetic sensitivity (Figure 1). Thus, laptop owners with higher aesthetic sensitivity scores put more laptop stickers on their laptops, and raters used number of stickers as a cue for aesthetic sensitivity. In other words, number of stickers "explained" a substantial amount of the accuracy effect for aesthetic sensitivity. Because such a modified lens model was significant for only one of 21 possible models (5 Big Five traits + 15 Big Five facets + narcissism), this result could be a false-positive one (Type-I error).

That university students reliably gauged other students' extraversion and open-mindedness based solely on their laptop stickers is remarkable and speaks to the predictive power that thin slices of information—or residual markers of behavior—can have on personality judgment. One potential drawback of this research is that we neglected to assess peer reports of laptop owners' personality

traits, which would add greater validity beyond self-reports. Another limitation is that—inspired by Gosling et al.'s (2002) seminal work—we used only eight raters; using a larger, more diverse sample—even if it results in planned missing data (where perceivers do not rate all targets, e.g., Webster & Campbell, 2022)—would have improved the generalizability of our results, and CC-SEMs can accommodate designs with missing data. We hope that future research into personality judgments based on people's objects, spaces, or residual behaviors consider using a CC-SEM approach over traditional methods.

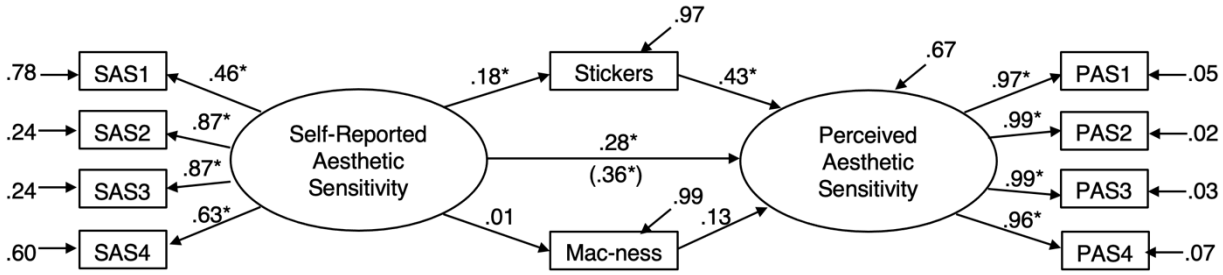


Figure 1. Results from the between-target level of a cross-classified structural equation model (CC-SEM) that examines log number of stickers as a cue for—or mediator of—the accuracy effect of the aesthetic sensitivity facet of open-mindedness. Standardized coefficients are shown; standardized coefficient of direct effect without cues or mediators appears in parentheses. Unstandardized coefficients and significant indirect effect via stickers are given in the main text. Based on 1,139 observations cross-classified across eight perceivers and 147 targets (laptops and their owners). SAS1–4: Self-reported aesthetic sensitivity items. PAS1–4: Perceived aesthetic sensitivity items. Mac-ness = probability that laptop is an Apple Macintosh (vs. PC; 0, .25, .50, .75, 1). Stickers = natural log of number of laptop stickers. *$p < .05$.

Preprint: https://psyarxiv.com/kz9wv

Data and Mplus code: https://osf.io/4ygca

References

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

Campbell, J. T., Turner, I. N., & Webster, G. D. (2022). Open laptops, open minds: Consensus and accuracy in Big Five personality perception from laptop stickers. *Journal of Research in Personality*, *96*, 104186. https://doi.org/10.1016/j.jrp.2021.104186

Claus, A. M., Arend, M. G., Burk, C. L., Kiefer, C., & Wiese, B. S. (2020). Cross-classified models in I/O psychology. *Journal of Vocational Behavior*, 103447. https://doi.org/10.1016/j.jvb.2020.103447

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*(3), 379–398. https://doi.org/10.1037/0022-3514.82.3.379

Harari, G. M., Graham, L. T., & Gosling, S. D. (2015). Personality impressions of world of Warcraft players based on their avatars and usernames: Consensus but no accuracy. *International Journal of Gaming and Computer-Mediated Simulations*, *7*(1), 58–73. https://doi.org/10.4018/IJGCMS.2015010104

Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the Dark Triad. *Psychological Assessment*, *22*(2), 420–432. https://doi.org/10.1037/a0019265

Nestler, S., & Back, M. D. (2013). Applications and extensions of the lens model to understand interpersonal judgments at zero acquaintance. *Current Directions in Psychological Science*, *22*(5), 374–379. https://doi.org/10.1177/0963721413486148

Nestler, S., & Back, M. D. (2017). Using cross-classified structural equation models to examine the accuracy of personality judgments. *Psychometrika*, *82*(2), 475–497. https://doi.org/10.1007/s11336-015-9485-6

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Webster, G. D., & Campbell, J. T. (2022). Personality perception in Game of Thrones: Character consensus and assumed similarity. *Psychology of Popular Media.* Advance online publication. https://doi.org/10.1037/ppm0000398

Comparing and Contrasting Rates of Firearm Homicides and Suicides

Rachel Christine Zambrano, Mateo Peñaherrera-Aguirre, & Aurelio José Figueredo

University of Arizona, United States

✉ rzambrano@arizona.edu

The clarification of the limited scope of the Dickey Amendment as only prohibiting funding for gun control advocacy in firearm violence research being funded by government agencies in the USA (Subbaraman, 2019) has spurred an influx of research on firearm homicide and suicide at many levels of government organization. Much of this research has focused on firearm homicide and firearm suicide as separate behaviors. Brunswikian theory provides a novel way to investigate the relationship between firearm homicide and suicide and identify if the two violent behavioral outcomes share any risk or protective factors.

The Brunswik Lens Model can be briefly summarized as a graphical representation of the way that living organisms integrate arrays of incoming sensory cues (inputs) to come to the perceptual decisions most correctly identifying events in the external environment and then select from arrays of alternative motoric means (outputs) to come to the behavioral decisions most effectively responding to the environmental challenges or opportunities indicated by incoming cues (Brunswik, 1952, 1955; Petrinovich, 1979). This model has been applied to problems of behavioral development, modeling not just external circumstances but any *strategically relevant* traits possessed by the developing organism (e.g., Figueredo et al., 2000), meaning traits that might affect the relative effectiveness of alternative means of achieving external objectives.

The correspondence between Brunswik's ecological views and contemporary evolutionary perspectives on cognition and behavior are not coincidental. Brunswik's endorsement of ecological psychology (some academics attribute the invention of this construct to Brunswik; Hammond, 2001), clearly indicates that the author supported a scientific understanding of ecological phenomena, wherein the environment referred to "the measurable characteristics of the objective surroundings of the organism rather than the psychological environment or life space" (Brunswik, reprint 2001, *p.* 300). Consequently, ecology corresponded to the sum total of all objective surroundings of an organism or taxa (Brunswik, reprint 2001). Brunswik's ecological psychology also operated as a prescriptive epistemological scaffold whereby the author suggested that environmental factors should be considered, from a theoretical and empirical lens, alongside the organism (Hammond, 2001). Furthermore, his view of probabilistic functionalism provided an in depth theoretical and methodological understanding of function or purpose in natural environments (Gigerenzer, 2001). This epistemological lens provides evolutionary researchers with a foundational framework for understanding interactions between organisms and environment.

Many Lens Models rely exclusively on individual-level traits to predict individual-level outcomes, but the science of behavioral ecology predicts that certain population-level phenomena (the *social ecology*) should also act as reliable and valid cues that might optimally guide individual behavior (e.g., Figueredo, Patch, & Gómez Ceballos, 2015). Thus, although Brunswik's claims related to

the interactions between organisms and environment, predominantly referring to individual-level process, it is theoretically feasible to generalize the author's views to higher levels of aggregation. For example, an individual observing the operational sex ratio of its population might use this as a cue indicating the degree of *local mate competition* (e.g., Gardner & Hardy, 2020) and use that information to adjust its mating and reproductive strategies. Similarly, an individual observing the external morbidity and mortality of its population might use this as a cue indicating its own expected morbidity and mortality in that same ecological setting (Ellis et al., 2009) and use that information to adjust its life history strategy from a faster to a slower speed or *vice versa*, as might be indicated by attaining an earlier or later age of puberty, age of sexual debut, or age at birth of the first offspring.

Under this light, biodemographic, macroecological, and epidemiological phenomena are predicted to have top-down influence individual-level phenotypes. The causal connections between these nested layers of organization are thought to operate via the organism's capacity to extract information from the environment in relation to the mean occurrence of particular events. Hence, the environment operates as a probabilistic landscape guiding the organism towards different phenotypic outcomes. Inspired by Brunswik's environmental perspectives, we hypothesized that ecological predictors would have different effects on firearm homicide and firearm suicide rates.

A recently published statistical model used various aggregate sociodemographic parameters of the 50 States of the USA to predict the aggregate suicide and homicide rates of the same entities (Zambrano et al., 2022). We incorporated many environmental factors that are not considered in "traditional research". Investigating social-biographical factors, such as, temperature, participation, parasite burden, psychopathology rate, firearm possession rate, and estimated IQ, allowed for us to establish a more externally valid model of firearm homicide and suicide. This analysis resulted in some unique results, such as local parasite burden strongly and positively predicting firearm homicide rate. In contrast, firearm possession rate, state psychopathology rate, GDP per capita, and estimated IQ were negatively related to firearm homicide. Firearm suicide was positively related to violent crime rate, firearm possession rate, poverty rate, and estimated IQ and negatively related to parasite burden and annual precipitation. These results display the wide range of alternative means that can predict the act of an individual committing homicide or suicide with a firearm. While access to firearms has been proposed as a variable that lowers the threshold for violence, we found that there were many other contextual factors that may also predict firearm homicide and suicide. Consistent with Brunswik's Lens Model, we found that the outcome of firearm homicide and suicide is a product of interaction between contextual and biological variables (physical setting, community ecology, and biological-psychological situation) to produce multiple interchangeable alternative means that can condition the likelihood of committing homicide or suicide with a firearm.

We apply these data to infer how individuals within those social ecologies might utilize such aggregate sociodemographic parameters to guide their individual behavioral decisions within those environmental contexts. At this point in the research program, we regrettably lack individual-level data to validate these inferences. However, because the external environment is being used as a source of cues by the individuals within it, Simpson's Paradox (1951) does not directly apply as we are not directly generalizing from aggregate to individual behavior. Employing Brunswik's concepts of ecological psychology and the Lens Model to firearm homicide and suicide allowed us to assess multiple points of contingency and their potential for convergence that alters an

individual's behavior. This theoretical framework facilitated investigating firearm homicide and suicide in a fruitful multidimensional manner that illuminated the components that condition these deadly behavioral outcomes.

References

Brunswik, E. (1952). The conceptual framework of psychology. In *International Encyclopedia of Unified Science*, vol. 1. University of Chicago: Chicago, IL.

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

Brunswik, E. (2001 reprint). Scope and aspects of the cognitive problem. In K. R. Hammond & T. R. Stewart (Eds.). *The essential Brunswik: Beginnings, explications, applications* (pp. 298–300). Oxford University Press.

Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk: The impact of harsh versus unpredictable environments on the evolution and development of life history strategies. *Human Nature*, *20*, 204–268.

Figueredo, A. J., Patch, E. A., & Gómez Ceballos, C. E. (2015). A Life History Approach to the Dynamics of Social Selection. In Zeigler-Hill, V., Welling, L. L. M., & Shackelford, T. K., (Eds.) *Evolutionary Perspectives on Social Psychology* (Chapter 28, pp. 363–372). New York, NY: Springer.

Figueredo, A. J., Sales, B. D., Becker, J. V., Russell, K., & Kaplan, M. (2000). A Brunswikian evolutionary-developmental model of adolescent sex offending. *Behavioral Sciences and the Law*, *18*, 309–329.

Gardner, A., & Hardy, I. C. W. (2020). Adjustment of sex allocation to co-foundress number and kinship under local mate competition: An inclusive-fitness analysis. *Journal of Evolutionary Biology*, *33*(12), 1806–1812.

Gigerenzer, G. (2001). Ideas in Exile: The Struggles of an Upright Man. In K. R. Hammond & T. R. Stewart (Eds.). *The essential Brunswik: Beginnings, explications, applications* (pp. 445–452). Oxford University Press.

Hammond, K. R. (2001). How Probabilistic Functionalism Advances Evolutionary Psychology. In K. R. Hammond & T. R. Stewart (Eds.). *The essential Brunswik: Beginnings, explications, applications* (pp. 434–440). Oxford University Press.

Petrinovich, L. (1979). Probabilistic functionalism: A conception of research method. *American Psychologist*, 34(5), 373–390.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society: Series B (methodological)*, 13(2), 238–241.

Subbaraman, N. (2019). United States to fund gun-violence research after 20-year freeze. *Nature*, 577, 12 (2020). https://doi.org/10.1038/d41586-019-03882-w

Zambrano, R., Peñaherrera-Aguirre, M., Figueredo, A. J., & Jacobs, W. J. (2022). Comparing and contrasting rates of firearm homicide and suicides. *Evolutionary Behavioral Sciences*. https://doi.org/10.1007/s40806-021-00311-y

THE BRUNSWIK SOCIETY

STAY CONNECTED

Visit the Brunswik Society website at https://brunswiksociety.org/.

To receive information about the Society's annual meeting and newsletter, email info@brunswiksociety.org.

Follow us on Twitter! @BrunswikThe