

V-Seminar – Summer Term 2026

Cognitive Biases and Behavioral Economics

Prof. C. Borgoni & Prof. F. Herweg

Course abstract

Inspired by findings from psychology regarding how an individual or her brain processes information and makes decisions, philosophers and economists have independently conceptualized these findings in their respective fields. The seminar will look at biases and heuristics used by individuals in decision-making from both a philosophical and an economic perspective. There will be discussions of broader, rather philosophical questions, e.g. to which extent the usage of heuristics is irrational and how the individual's mind can carry them out, as well as more applied – economic – questions, e.g. how do these biases affect tradeoffs between short-run benefits and long-run goals (savings for retirement).

This year's edition also includes a dedicated topic on biases in generative AI, as well as a group activity in which students will design a simple empirical study.

Organizational Issues

Dates and deadlines

- **Introduction and Organization session: April 16th, 2026, 10:00 (ZOOM)**
- **Deadline for Registration:** April 26th, 2026
- **Submission of Preference List:** Each accepted student should email Fabian.herweg@uni-bayreuth.de naming her/his preferred three topics by April 26th.
- **Assignment of Topics:** April 27th
- **Seminar:** July 2nd – 3rd, 2026 (approx. from 9:30am until 4:00pm), tba
- **Submission of a short summary of basic literature** (upload to eLearning): June 26th
- **Submission of term papers** (uploaded to eLearning): July 31st, 2026

Target group

Advanced bachelor's students from P&E, Economics, and IWE.

Enrollment and seminar places

At most 20 students are accepted for the seminar. Students can enroll for the course on the *CampusOnline* platform. The procedure will be explained in the "Introduction and Organization session".

Course language

Presentations should be held, and term papers should be written in English.

Assessment

Each student has to give a presentation, actively participate in the discussion, and write a **term paper**. You must also write a **1-page summary (in total)** of the basic literature. On the first day of the seminar, there will be a quick writing quiz on the basic literature. The term paper should be 8-9 pages long, not counting appendices. Each presentation should be approximately 30 minutes, followed by 20 minutes of discussion. If we have more than 10 students enrolled, presentations will be done in groups (of two students).

- **Philosophy & Economics:** 6 cp, V3.
- **Economics:** 5cp, Seminar “Institution und Governance“, „Mikrotherotisches Seminar“.
- **Internationale Wirtschaft und Entwicklung:** Spezialisierung IGME, VET.

Seminar Topics

The Basics: Cognitive Biases, Heuristics, and Implicit Biases (group discussion)

- *Objective:* To be acquainted with the basic literature on heuristics and implicit biases that forms the basis of the course. Everyone is required to read the literature below and write 1-page summary of every reading.
 - Tversky, A. & Kahneman, D. (1974). Judgement under Uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.
 - Tversky, A. & Kahneman, D. (1983). Extensional vs. Intuitive reasoning: The Conjunction Fallacy in Probability Judgement. *Psychological Review*, 90. 293-315.
 - Nosek, B; Hawkins, C and Frazier, R. (2011), “Implicit social cognition: From measures to mechanisms”, *Trends Cognitive Science* 15(4): 152–159.

The 10 Topics of the Seminar

1. The Nature of Implicit Attitudes

- **Question for presentation (Q3):** What is the nature of implicit biases (i.e., what type of mental state they are)?

Short explanation: In philosophy of mind, we distinguish between various types of mental states: beliefs, perceptions, desires, intentions, emotions, etc. Each type of mental state has different constitutive aspects that determine their function in our mental life and also their patterns of behavior and activation. Beliefs, for example, are normally taken to be responsive to reasons, and to aim at being true. The question of this topic is whether implicit biases can be correctly classified as

beliefs. Answering this question has consequences to the way one ascribes responsibility and rationality to the agent in regards to the presence of such biases.

- Primary readings:
 - Gendler, T., (2008), “Alief and Belief”, *Journal of Philosophy*, 105(10): 634–63.
 - Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, 131(4), 1753-1794.
- Complementary readings:
 - Bendana, J. (2021), “Implicit Attitudes Are (Probably) Beliefs”, in: Borgoni, C. Kindermann, D. and Onofri, A. *The Fragmented Mind*, Oxford: OUP, 2021.

2. Epistemic costs and benefits of biases and heuristics

- **Question for presentation (Q2):** Do heuristics lead us to (mostly) correct or incorrect decisions and judgments?

Short explanation: The accepted view in philosophy is that a judgment is correct when it is true. Another aspect of correctness is that it is well justified by reasons. Perhaps one would like to say that a correct judgment is the one whose truth is reached in the ‘right’ way. Are heuristics good routes to the truth? Are there epistemic costs in trying to avoid heuristics that lead us to false judgments?

Primary reading:

- Gendler, Tamar (2011) “On the epistemic costs of implicit bias”, *Philosophical Studies* 156: 33–63.
PS: Be critical regarding the way the author describes the calculus of probability of a black person being a criminal given the race rate in prison population.

Complementary reading:

- Borgoni, C. (2022), “Philosophy of Mind after Implicit Biases”, *The Political Turn in Analytic Philosophy: Reflections on Social Injustice and Oppression*, Berlin: De Gruyter, 2022.

3. Moral Responsibility

- **Question for presentation (Q4):** Are we responsible for the outputs of our biases and heuristics?

Short explanation: Biases, including implicit biases, and heuristics belong to the part of our mind that is not easily accessible via introspection or via reasoning. Many would say that they are part of our sub-personal system. If so, if we apparently lack full control or awareness of such mechanisms, can we be held responsible for the decisions and judgments that result from them? Can we be blamed for being biased in this way?

- Primary reading:
 - Holroyd, Jules (2012), “Responsibility for implicit bias”, *Journal of Social Philosophy* 43 (3), pp. 274-306.
- Complementary readings:
 - Levy, Neil (2014) “Consciousness, Implicit Attitudes and Moral Responsibility”, *Nous* 48:1, pp. 21-40.

4. Rationality

- **Question for presentation (Q5):** Do heuristics make us rational or irrational?

Short explanation: One of the great impacts of the discovery of biases and heuristics is that our decisions and judgments very often differ from the standards established by logics and other theoretical approaches to reasoning. Does that make us irrational? Should logics or nature establish the standards of what is rational?

- Primary reading:
 - Rich, Patricia. “Axiomatic and Ecological Rationality: choosing costs and benefits”. *Erasmus Journal for Philosophy and Economics* (9): 90-122.
- Complementary reading:
 - Smith, V. L. “Constructivist and Ecological Rationality in Economics”, *The American Economic Review*, 445-508.

5. Biases in AI systems

- **Question for presentation (Q5):** In which sense do generative AI preserve and amplify human heuristics and implicit biases?

Short explanation: the objective is to understand how AI may amplify, rather than correct, morally and epistemically deviant human biases, and reflect on how counter such effects.

- Primary reading:
 - Zhou, M., Abhishek, V., Dardenger, T., Kim, J., & Srinivasan, K. (2024). Bias in generative AI. arXiv. <https://doi.org/10.48550/arXiv.2403.02726> (pre-print)
 - Motoki, F. Y. S., Pinho Neto, V., & Rodrigues, V. (2025). *Assessing political bias and value misalignment in generative artificial intelligence*. *Journal of Economic Behavior & Organization*, 234, 106904. <https://doi.org/10.1016/j.jebo.2025.106904>
- Complementary reading:
 - Ananya. 2024. “AI Image Generators Often Give Racist and Sexist Results: Can They Be Fixed?” *Nature* 627 (8005): 722–725 (non-academic piece).

6. Two-Modules View of the Brain

- **Question for presentation (Q7):** What are the different objectives of the doer and the planner?

There is an old literature in psychology in which the brain is viewed as comprising of two modules: an emotional and a cognitive module (Schneider and Shiffrin, 1977). This idea has been used by economists to model the behavior of individuals as the outcome of a competition between the two modules (game-theoretic equilibrium).

- Primary reading:
 - Shefrin, H.M., & Thaler, R.H. (1988). The Behavioral Life-Cycle Hypothesis. *Economic Inquiry*; 26(4), 609- 643.
- Complementary reading:
 - Fudenberg, D., & Levine, D. K. (2006). A dual-self model of impulse control. *American Economic Review*, 96(5), 1449-1476.
 - Evans, J. “How Many dual-process theories do we need? One, two, or many?”, in: *In Two Minds: Dual Processes and Beyond*, Oxford: OUP: 33-54.
 - Bernheim, B. D., & Rangel, A. (2004). Addiction and cue-triggered decision processes. *American Economic Review*, 94(5), 1558-1590.

7. Mental Accounting

- **Question for presentation (Q8):** What is the psychological principle behind the idea of mental accounting?

The objective is to discuss a theory that takes into account that individuals do not take all interrelations into account when making decisions; i.e., individuals group decisions in certain categories and make tradeoffs only within a category.

- Primary reading:
 - Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, 4(3), 199-214.
- Complementary reading:
 - Prelec, D., & Loewenstein, G. (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science*, 17(1), 4-28.
 - Köszegi, B. and F. Matějka (2018). Choice Simplification: A Theory of Mental Budgeting and Naive Diversification. *Quarterly Journal of Economics*, 135(2), 1153–1207.

8. Libertarian Paternalism – A Critique

- **Question for presentation (Q9):** Should (benevolent) authorities steer peoples’ choices in the right directions?

Evidence from psychology, marketing, and behavioral economics documents that people often do not make rational choices and that choices can (easily) be manipulated. Based on this evidence, some economists argue that authorities should intervene so that people make “better” decisions. Ideally, this interventions do not restrict choices so that smart decision makers are not negatively affected.

This is the main idea of libertarian paternalism. A prominent notion of this kind of intervention is *nudging*.

- Primary reading:
 - Gigerenza, G. (2015). On the Supposed Evidence for Libertarian Paternalism, *Review of Philosophical Psychology*, 6, 361-383.
- Complementary reading:
 - Thaler, R. H., and C. R. Sunstein (2003). Libertarian Paternalism, *American Economic Review P&P*, 93 (2), 175-179.

9. The Positive and the Normative Value of Canonical Choice Axioms

- **Question for presentation (Q9):** When do decision-makers consider it a mistake to violate canonical axioms of choice, and what does this tell us about the normative attractiveness of these axioms?

Economic analysis is based on the idea of axiomatic rationality. A decision is considered rational if it obeys certain axioms. For choices under risk it is well-documented that decisions often violate canonical axioms. An important question therefore is whether decision-makers who violate a particular axiom consider this violation as a mistake. The goal of this topic is to gain a better understanding of when choices that violate canonical axioms are mistakes and thus considered as non-rational choices by the decision-maker herself.

Primary reading:

- Nielsen, K & J. Rehbeck (2022). When Choices Are Mistakes, *American Economic Review*, 112(7), 2237-2268. Complementary reading:
- Herweg, F., S. Hippel, D. Müller & F. Römeis (2024). Axiom Preferences and Choice Mistakes under Risk, Working Paper, University of Wuerzburg.

10. Recasting Welfare Economics

- **Question for presentation (Q10):** How can we conduct welfare analysis based on choices if preferences are not stable, and thus choices do not necessarily reflect “true” preferences?

Non-rational decisions, which do not reflect stable preferences, are a challenge for standard welfare economics (based on revealed preferences). Nevertheless, without presumptions about what good choices are, welfare analysis can be based only on observed choices. Does this imply that we – as economist – have to give up on welfare analysis in the age of behavioral economics?

- Primary reading:
 - Sunstein, C.R. (2020). Behavioral welfare Economics, *Journal of Benefit-Cost Analysis*, 11(2), 196 - 220.
- Complementary reading:

- Kolmar, M. (2026). The Hidden Objectivism of Revealed-Preference Welfare
- Economics *CEsifo Working Paper*, 12519.
- Bernheim, B.D. & A. Rangel (2007). Toward Choice-Theoretic Foundations for Behavioral Welfare Economics, *American Economic Review P&P*, 97(2),464-470.