Building Computational Method Skills in First-Year Economics: Programming, MathStat & Gamification

September 15, 2025

Dr. Johannes Bleher University of Hohenheim

https://johannes-bleher.de

johannes.bleher@uni-hohenheim.de

Motivation

- I would like to know what you are doing.
- Economics students need strong computational and quantitative skills.
- Traditional lectures (math/stat only) risk losing student motivation.
- What I try to combine:
  1. A feeling for Data (Spreadsheets/Excel)
  2. Programming (R, Python, Stata - AIDAHO Jupyterhub)
  3. Math/Stat foundations
  4. Gamified learning via web apps
  5. Integration of AI

Teaching Context (cont.)

- Since April 2025: Akademischer Rat ($\sim$ Assist. Prof) at University of Hohenheim.
- Took over statistical training for the first-/second-year students of the Faculty of Business, Economics and Social Sciences ($\sim$ 900 students)
  - **Past summer term:** *Intro to Statistical Data Analysis.*
  - **Coming winter term:** *Sample-based Data Analysis.*
- Previous courses I taught in Hohenheim:
  - *Tools for AI & Data Science: Intro to R, Python and SQL*
  - *Introduction to Data Science with R and RStudio*
  - *Applied Data Science Lab*

Teaching Context (cont.)

- Teaching Award 2021 of the University of Tübingen (with Joachim Grammig and Thomas Dimpfl) for integrating programming and games into the first-year Math/Stats economics curriculum.
  - Implementation of a JupyterHub on bare-metal Kubernetes setup
  - Integrated R and Python in Math Prep class
  - Gamification of linear algebra
  - Virtual PC-Lab – virtual machines on demand in the (openstack) bwCloud

  See Bleher and Dimpfl (2021) for documentation (in German).

Teaching Context (cont.)

- **Current goal:** Build on these experiences and integrate proven didactic techniques and digital resources into the Hohenheim curriculum.
  - Use the AIDAHO-JupyterHub or bwJupyterHub for an easy start to programming.
  - Use the GRAIDAHO automated code assessment platform for self-testing.
  - Integrate WebGames in teaching.

General Research Context (cont.)

- **Spacing / distributed practice** improves long-term retention and transfer (Cepeda et al., 2006, 2008).
- **Desirable difficulties** – self-testing, generation, varied retrieval, effortful processing – deepen learning (Bjork, 1994; Bjork and Bjork, 2011).
- **Interleaving / randomization** outperforms blocked practice for discrimination and transfer (Shea and Morgan, 1979; Rohrer and Taylor, 2007).

Popularized also in the audiobook Polk (2018).

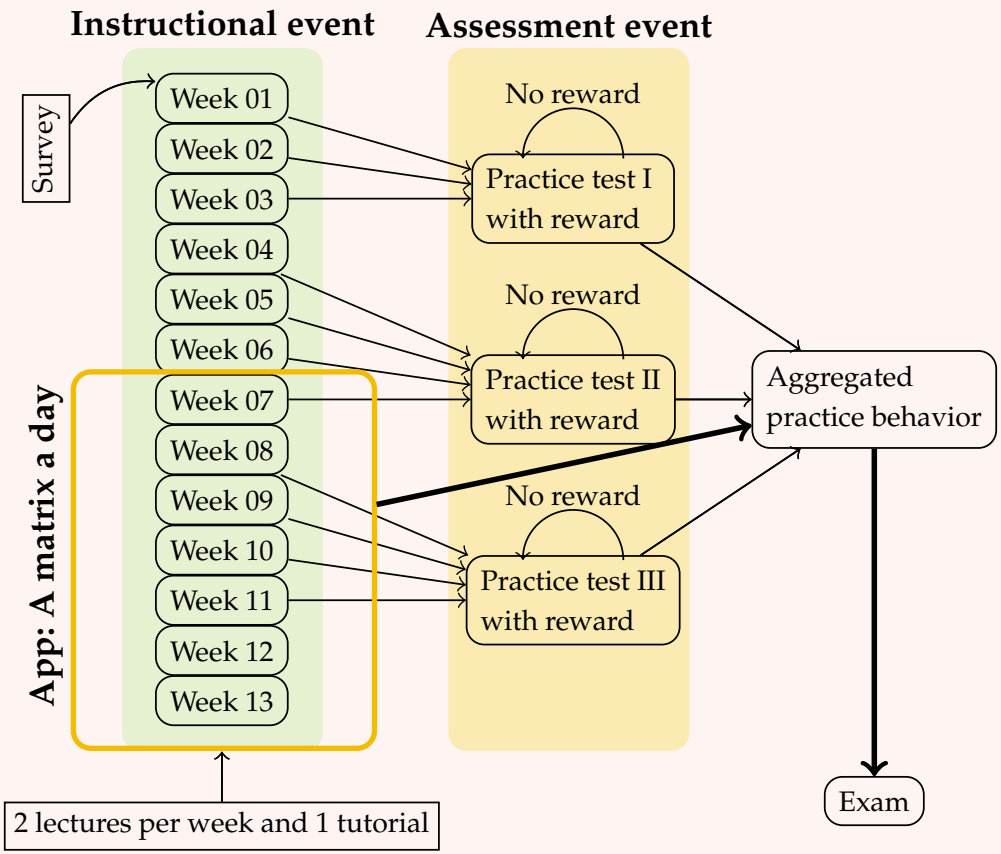
General Research Context (cont.)

- **Gamification** uses points, badges, leaderboards to increase engagement.
  - Reviews show mixed but overall positive effects on motivation and learning (Hamari et al., 2014; Seaborn and Fels, 2015; Subhash and Cudney, 2018).
  - Empirical classroom studies highlight links between challenge, flow and deeper learning (Hamari et al., 2016).
  - Causal evidence: Randomized controlled trial demonstrates improved exam performance through gamification in higher education (Denny et al., 2018).

Personal Research Context

Benefits of Additional Online Practice Opportunities in Higher Education

Schwerter et al. (2022)



Sample and Context

- Economics/business administration students (N=188, 59% female)
- 1st semester gateway course
- Practice tests: Content from the previous 3-to-4 weeks; up to +2pts in exam for 1st trial
- Immediate corrective feedback
- Explanatory variables:
  - Practice test attempts
  - Practice test performance
  - App submissions
  - App performance
- Outcome: Points in final exam

Results (cont.)

	Dep. variable: Standardized points on final exam			
	Practice vars only	PDSR LASSO	PDSR Rand. For.	PDSR XGBoost
Constant	-2.401***	-0.824	-0.078	-0.195
	(0.328)	(0.847)	(1.256)	(0.753)
Practice test attempts	0.226***	0.215***	0.203***	0.205***
	(0.074)	(0.068)	(0.065)	(0.064)
Practice test performance	0.022***	0.010**	0.010**	0.010**
	(0.004)	(0.004)	(0.004)	(0.004)
App submissions	0.007	0.006	0.001	0.004
	(0.012)	(0.008)	(0.009)	(0.009)
App performance	0.004**	0.004**	0.005***	0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
Additional controls	No	Yes	Yes	Yes
Observations	188	188	188	188
Adjusted R^2	0.213	0.446	0.464	0.410

PDSR: Post-double selection regression ⇒ Selecting important variables for the dependent variable and the four practice variables using LASSO, RF, and XGBoost.

Results (cont.)

Variable importance

	Variables	Importance
1	HS GPA	0.42
2	Practice test performance	0.22
3	Math self-concept	0.18
4	HS math grade	0.16
5	Practice tests attempts	0.11
6	App submissions	0.09
7	App performance	0.08

Prediction rules

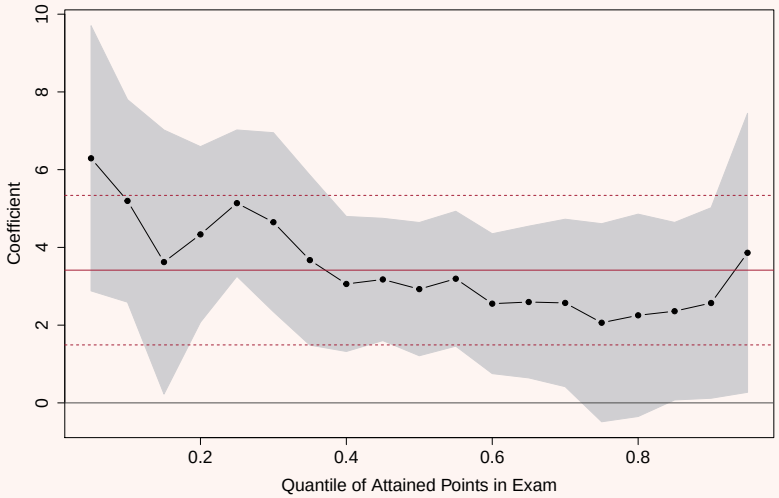
Rules	Coefficient
$SC_{Math} \leq 3.25$ & App submissions ≤ 3	-0.37
HS GPA ≤ 3.7 & App performance ≤ 45.71	-0.35
PT attempts ≤ 3 & HS GPA ≤ 3.7	-0.34
HS GPA > 3.3 & HS math grade > 2.6	0.30
HS GPA > 4.1 & PT performance > 65.33	0.28
HS GPA > 3.6 & HS math grade > 2.6	-0.20
HS GPA > 3.7 & PT performance > 78.03	0.20
PT performance > 56.67 & $SC_{Math} > 2.67$	0.17
PT performance > 56.67 & HS math grade > 3	0.16
HS GPA > 3.3 & $SC_{Math} > 2.25$	0.16
PT attempts ≤ 3 & PT performance ≤ 72.41	-0.14

PT = Practice test

*Fokkema, M., & Strobl, C. (2020). Fitting Prediction Rule Ensembles to Psychological Research Data: An Introduction and Tutorial. *Psychological Methods*, 25(5), 636–652. https://doi.org/10.1037/met0000256

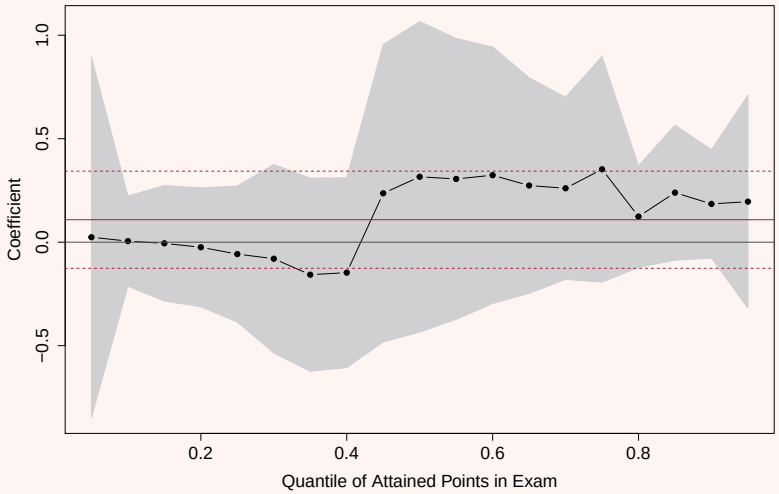
Results (cont.)

Figure: Practice test attempts



- Students with fewer exam points would have benefitted more than those who attained more points

Figure: App submissions



- Somewhat constant null effect
- Population-regression: Good students would have benefitted more from practice in matrix algebra

Results (cont.)

Figure: Practice test performance

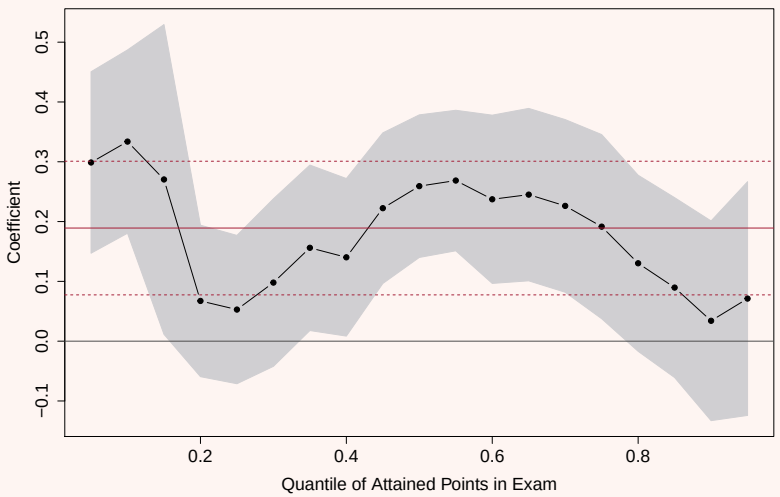
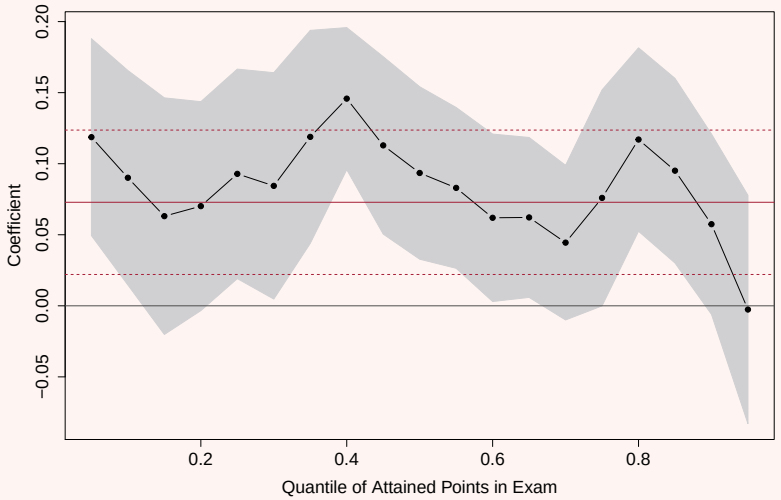


Figure: App performance



- Students with fewer and mediocre exam points would have benefitted more from practicing to mastery

Conclusion (cont.)

💡 **Findings:**

- One practice test participation improved final exam scores by ~ 5 points (OLS); ~ 2.5 points with rich covariates.
- Strongest gains for lower prior achievers.
- Course-embedded tests outperformed daily app-based self-testing.
- Learning until achieving good performance in self-testing pays off.
- For some students: The game app seemed to form a habit (even after winners were announced continued submissions).

Infeasible in 2025 (GenAI)

- **LLMs solve items**
- **Ubiquitous access:** Phone camera $\rightarrow$ instant solution.
- **Proctoring limits:** Home environments, privacy law, and false positives make reliable exam supervision unrealistic.
- **Fairness & legal risk:** Reward points tied to unverifiable work introduce inequity and contestability.

Bottom line: Remote, *rewarded* ILIAS practice tests can no longer assure authorship *at scale* $\Rightarrow$ incentives distort.

But: Other incentives (through gamification) might still work.

Cheat Sheets

Prior A 180 pages collection of formulas allowed in exam.

Now 1 handwritten DIN A4 paper sheet. But: Students can negotiate the number of DIN A4 sheets by handing in their sheets 6 weeks prior to the exam.

- ⭐ Students *space* learning, *challenge* themselves: 'What is really important?', *interleave* materials.
- ⭐ Statistical assessment of handed-in sheets with (subjective) compression rate and coverage rate.
- ⭐ Students hated ex-ante, lauded ex-post
- 🧭 TODO: Automate with multi-modal models.

Web Games

Prior No games used in teaching.

Now Used the game 'Guess the Correlation' in past term. Developing a game for probability calculations (Bayes law) with daily 'crime' cases and a leader board *Conditional Statistics Investigation Hohenheim (CSI-HOH)*

- ⭐ Students *space* learning, *challenge* themselves through self-testing.
- ⭐ Web development and scenario generation with LLMs is fast.
- 🧭 TODO: Interactive features for students to write scenarios via code.

Computational Methods and Programming

Prior Stata sessions with 3 sessions - voluntary extra points for an assignment.

Now Integrating Excel, Stata, R and Python into the lecture, as well as regular exercise sheets and the exam. Extra points are also achievable for an assignment.

- ⭐ Students *challenge* themselves through implementing program code, self-testing, and interleaving theoretical considerations with concrete data.
- 🧭 TODO: Usage of self-testing platform GRAIDAHO.

WebApp Framework for Games and other Apps

- Backend: **Tornado (Python)** for background tasks.
- Frontend: **React + Tailwind (JavaScript)** for UI.
- Containerized architecture for easy deployment and replicability. Connected through an NGINX container.
- LLM-endpoint (on-premise, OpenAI, Google, ...)
- Features:
  - Score Boards.
  - Scalability (runs on Hohenheim servers).
  - Interaction with LLMs is used to vary the scenarios.
- Examples:
  - A Matrix A Day Keeps Dr. Dimpfl Away.
  - The Marriage Simulator.
  - CSI Hohenheim (coming forth).

What do you do ?

- **How is undergraduate math and statistics taught at your university?**
- **How do you incentivize your students to space, challenge and randomize their learning?**
- **Are you using elements of gamification?**
- **Do you integrate computational methods?**

…

References (cont.)

Bjork, E. L. and Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In Gernsbacher, M. A., Pew, R. W., Hough, L. M., and Pomerantz, J. R., editors, *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*, pages 56–64. Worth Publishers, New York, NY.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In Metcalfe, J. and Shimamura, A. P., editors, *Metacognition: Knowing about knowing*, pages 185–205. MIT Press, Cambridge, MA.

References (cont.)

Bleher, J. and Dimpfl, T. (2021). *Computergestützte Methodenkompetenzvermittlung für Studienanfänger in den wirtschaftswissenschaftlichen Studiengängen*. Number 17 in Tübinger Beiträge zur Hochschuldidaktik. Universität Tübingen. TBHD (17, 2021, 1).

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3):354–380.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., and Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science*, 19(11):1095–1102.

References (cont.)

Denny, P., McDonald, F., Empson, R., Kelly, P., and Petersen, A. (2018). Empirical support for a causal relationship between gamification and learning outcomes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 311:1–311:13.

Hamari, J., Koivisto, J., and Sarsa, H. (2014). Does gamification work? – a literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS)*, pages 3025–3034.

References (cont.)

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., and Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54:170–179.

Polk, T. A. (2018). The learning brain (audiobook). Audible / The Great Courses.

Rohrer, D. and Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, 35:481–498.

Schwerter, J., Dimpfl, T., Bleher, J., and Murayama, K. (2022). Benefits of additional online practice opportunities in higher education. *Internet and Higher Education*, 53:100834.

References (cont.)

Seaborn, K. and Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, 74:14–31.

Shea, J. B. and Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory*, 5(2):179–187.

Subhash, S. and Cudney, E. A. (2018). Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior*, 87:192–206.