Decreasing Differences in Expert Advice: Evidence from Chess Players

Elias Bouacida[1] Renaud Foucart[2] Maya Jalloul[2]

[1]University Paris 8, France

[2]Lancaster University, United Kingdom

EEA-ESEM, Rotterdam, August 2024

Decreasing Differences

The higher your ability in a task, the less you benefit from advice about it

The Great Equalizer

- Imagine you are an expert at something: a doctor, an economist, a driver, or, in our case, a chess player
- Your job is to solve problems and, for each of them, you can access the advice a (top) external adviser
- If productivity $f(\textit{ability}, \textit{advice})$ displays *decreasing differences*, then current evolution in information technology should lead to lower inequality

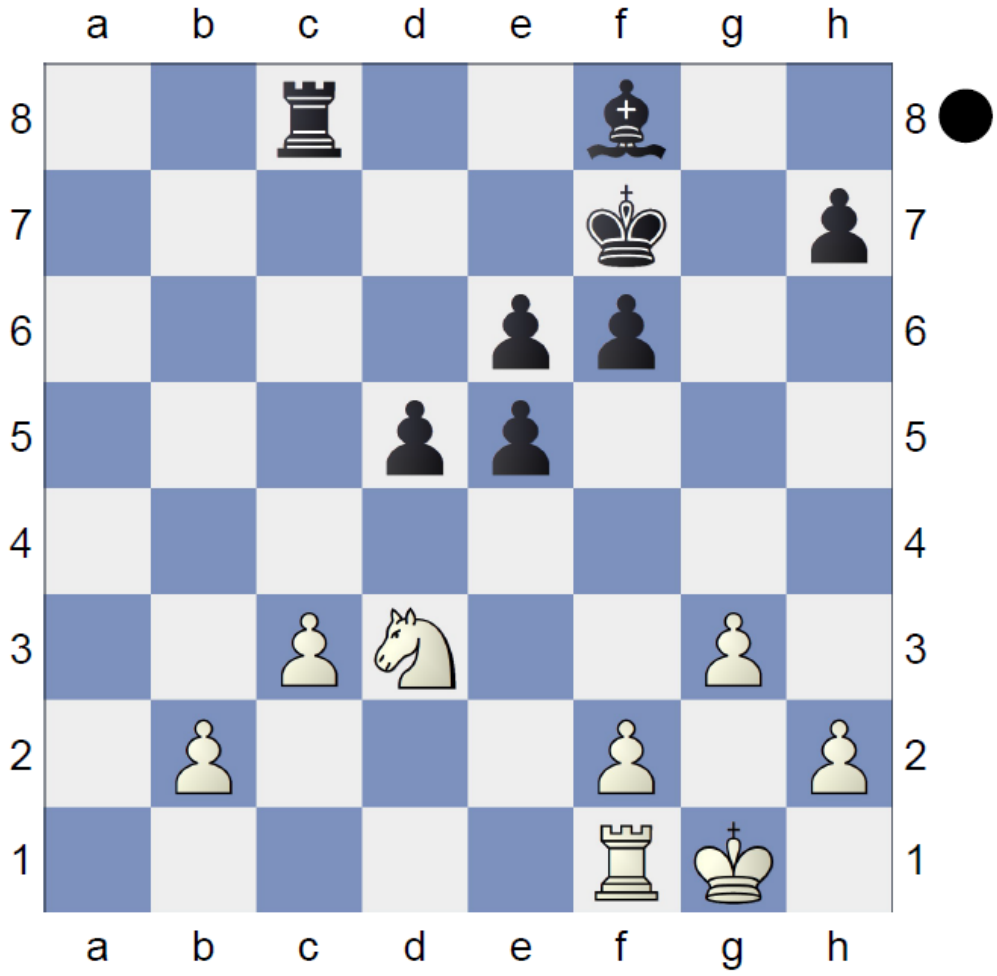
Mixed Evidence

- Evidence of decreasing differences in teams in factories (Hamilton *et al.*, 2003; Adhvaryu *et al.*, 2020), at university (Fischer *et al.*, 2023), in the US labour market (Herkenhoff *et al.*, 2024).
- Also from AI on lawyers (Choi and Schwarcz, 2023), programmers (Peng *et al.*, 2023), writers (Noy and Zhang, 2023), customer support (Brynjolfsson *et al.*, 2023), and consultants (Dell'Acqua *et al.*, 2023)
- But Agarwal *et al.* (2023) finds that radiologists often fail to incorporate uncertain advice optimally, and Otis *et al.* (2023) that, among Kenyan entrepreneurs, advice increases the performance of high performers but actually hurts low performers

This Paper

- We ask chess players to look at chess *positions* and evaluate the *pawn advantage*
- Subjects are competing in chess tournaments in Lebanon
- Advisers are an International Master (Elo rating of 2,335, among 6,000 best in the world) and a regular player playing for fun (no rating)

A Position



رقم الوضع Position Number	Part 1 - الجزء الأول			
	-2.4	-0.7	+0.7	+2.4
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Treatments

- We ask chess players to look at chess *positions* and evaluate the *pawn advantage*
- After a round of 10 evaluations, we show them the answers of one of the advisers
- In one treatment, we tell them the rating of the adviser, in the other, we don't

Known Adviser

A1i

We will now provide you with some additional information about the ten positions.

We have asked **a player with a rating of 2335** to evaluate the ten games in the same conditions as you. You can find their prediction in the table below.

Looking back at your own evaluation in **Round 1, Part 1** on the Response Sheet, please complete **Round 1, Part 2**. You are free to change or keep your previous predictions based on the information on this sheet.

Advice

رقم الوضع Position Number	التفوق Pawn advantage
1	-2.4
2	-0.7
3	-0.7
4	+2.4
5	+0.7
6	-0.7
7	+2.4
8	-0.7
9	-0.7
10	+2.4

You have a total of 4 minutes to complete this part.

لديك 4 دقائق لإكمال هذا الجزء.

Answer Sheet

Round 1 – الجولة الاولى:

رقم الوضع Position Number	Part 1 - الجزء الأول				Part 2 - الجزء الثاني			
	-2.4	-0.7	+0.7	+2.4	-2.4	-0.7	+0.7	+2.4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

The experiment

- Pre-registered incentivized lab-in-the-field experiment during chess tournaments in Lebanon[1]
- 102 players, 54 of them with a formal Elo rating, asked to evaluate chess positions Elo distribution
- 20 positions in total, each from a past game from Mega Database 2023, split in two “rounds” of 10
- Main results based on a binary outcome of whether the evaluation is correct or not

[1] https://aspredicted.org/124_MSY

Decreasing Differences

- f: production function
- H, L: higher- or lower-quality advice
- h, l: higher- or lower-ability subject

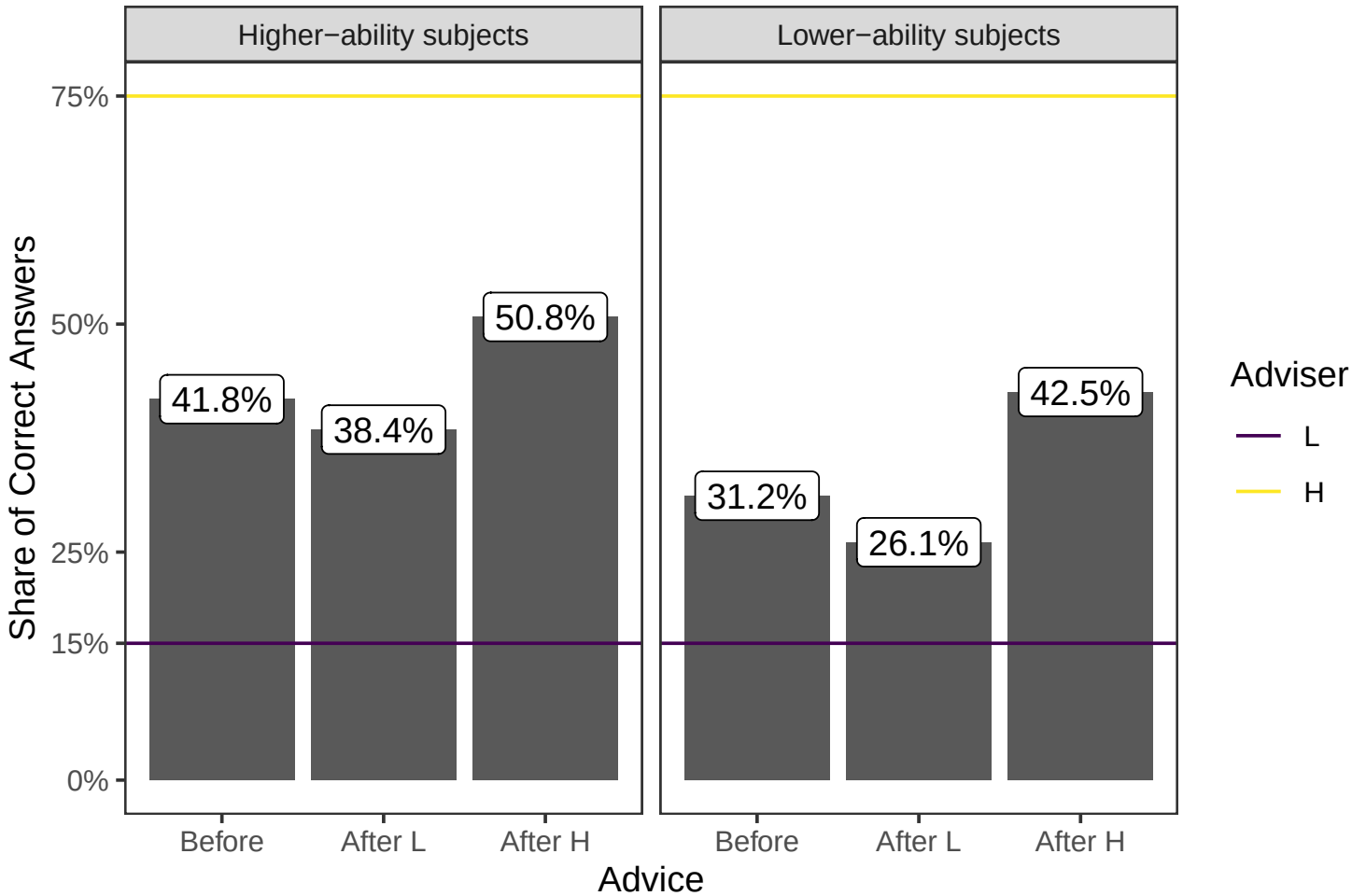
Decreasing differences

$$f(l,H) - f(l,L) > f(h,H) - f(h,L)$$

Negative vs Positive Assortative Matching

$$f(l,H) + f(h,L) > f(h,H) + f(l,L)$$

The Big Picture



By Treatment

NAM vs PAM

We compare the share of correct answers when *l* (resp. *h*) subjects receiving *H* (resp. *L*) advice (Negative Assortative Matching, NAM) and Positive Assortative Matching (PAM)

Treatment	**NAM**	**PAM**	**p-value**[1]
Known adviser	42.6%	41.3%	0.711
Unknown adviser	38.1%	35.2%	0.384

[1] p-value of the Welch Two Sample two-sided t-test of equality

Full pre-registered analysis

When Subjects Disagree with Advice

			Among those who disagree before[2]			
Subjects	Treatment	Disagree[1]	Keep	Follow	Closer	Further
Lower-ability	Know H	66.5	46.8	39.2	12.3	1.8
	Know L	72.2	78.8	14.0	7.3	0.0
	Unknown	70.6	64.3	27.3	6.3	2.2
Higher-ability	Know H	58.2	52.5	42.5	2.5	2.5
	Know L	73.2	83.9	9.5	4.0	2.5
	Unknown	67.0	79.7	15.3	3.3	1.7

We remove from this table the missing answers because we have no distance from the answer for them. We therefore slightly underestimate the disagreement percentage before receiving the advice.

[1] Percentage of different pre-advice answers with the adviser.

[2] Percentage of kept or changed answer (following, getting closer, or further away from advice) conditional on pre-advice answer being different from adviser's.

Table: What happens when subject disagree with the advice given?

Keeping by Treatment

How much is left on the table?

- We measure the difference between payoffs in the experiment and simple heuristics
- Consider subjects learning to maximize their share of correct answers by always following or ignore a certain quality of advice (*L*, *H* and unknown)
- How much do they leave on the table by ignoring advice?

Probability Heuristic

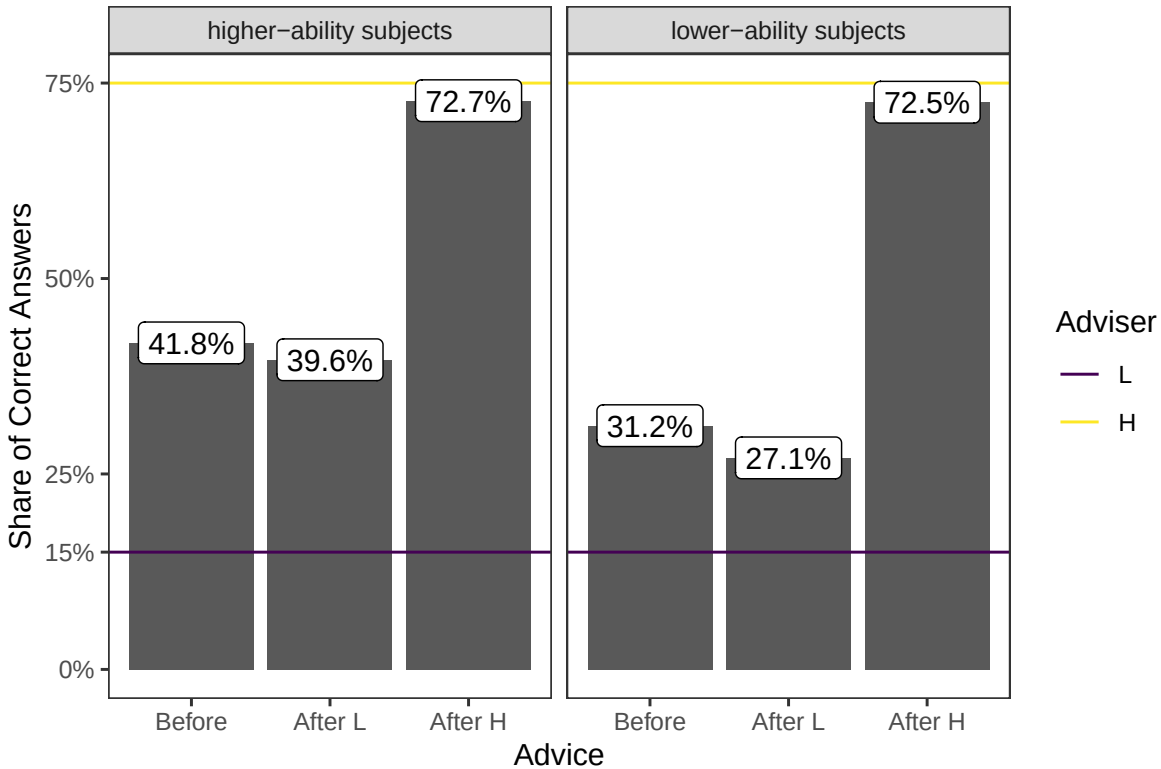


Figure: Share of correct answers following the probability heuristic.

Heuristics

Elo First-best

Loss After H-advice

	Subjects		
Treatment	Lower-ability	Higher-ability	P-value[1]
Unknown	32.4	23.0	0.056
Known	27.7	21.1	0.134
All	30.0	22.2	0.014

[1] P-value of the two-sided two sample t-test of the difference between *h* and *l* subjects.

Table: How much better subjects would have been following the heuristics after receiving high-quality advice - in percentage points?

The not-so-great equalizer

- A lab-in-the-field experiment in which we offer advice to subjects experts in their field
- No evidence of decreasing differences
- What we would love to know: why do lower-ability subjects leave more money on the table than higher-ability subjects? Preference for control? Overconfidence?

References I

Adhvaryu, A., Bassi, V., Nyshadham, A. and Tamayo, J. A. (2020). *No line left behind: Assortative matching inside the firm*. Tech. rep., National Bureau of Economic Research.

Agarwal, N., Moehring, A., Rajpurkar, P. and Salz, T. (2023). *Combining human expertise with artificial intelligence: Experimental evidence from radiology*. Tech. rep., National Bureau of Economic Research.

Brynjolfsson, E., Li, D. and Raymond, L. R. (2023). *Generative AI at work*. Tech. rep., National Bureau of Economic Research.

Choi, J. H. and Schwarcz, D. (2023). Ai assistance in legal analysis: An empirical study. *Available at SSRN 4539836*.

References II

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F. and Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of ai on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (24-013).

Fischer, M., Rilke, R. M. and Yurtoglu, B. B. (2023). When, and why, do teams benefit from self-selection? *Experimental Economics*, pp. 1–26.

Hamilton, B. H., Nickerson, J. A. and Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of political Economy*, **111** (3), 465–497.

References III

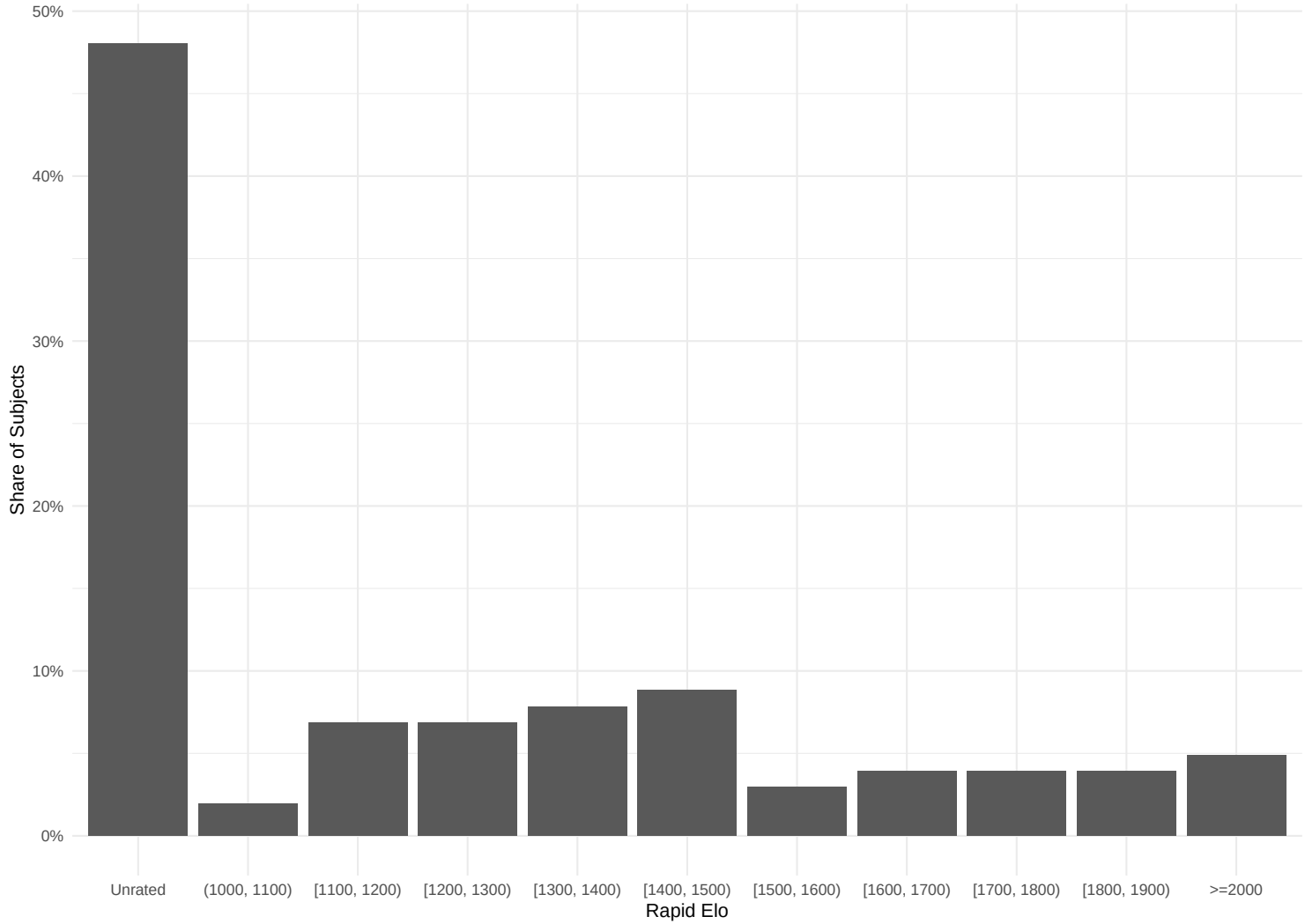
Herkenhoff, K., Lise, J., Menzio, G. and Phillips, G. M. (2024). Production and learning in teams. *Econometrica*.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Available at SSRN 4375283*.

Otis, N., Clarke, R. P., Delecourt, S., Holtz, D. and Koning, R. (2023). The uneven impact of generative ai on entrepreneurial performance. *Available at SSRN 4671369*.

Peng, S., Kalliamvakou, E., Cihon, P. and Demirer, M. (2023). The impact of ai on developer productivity: Evidence from github copilot. *arXiv preprint arXiv:2302.06590*.

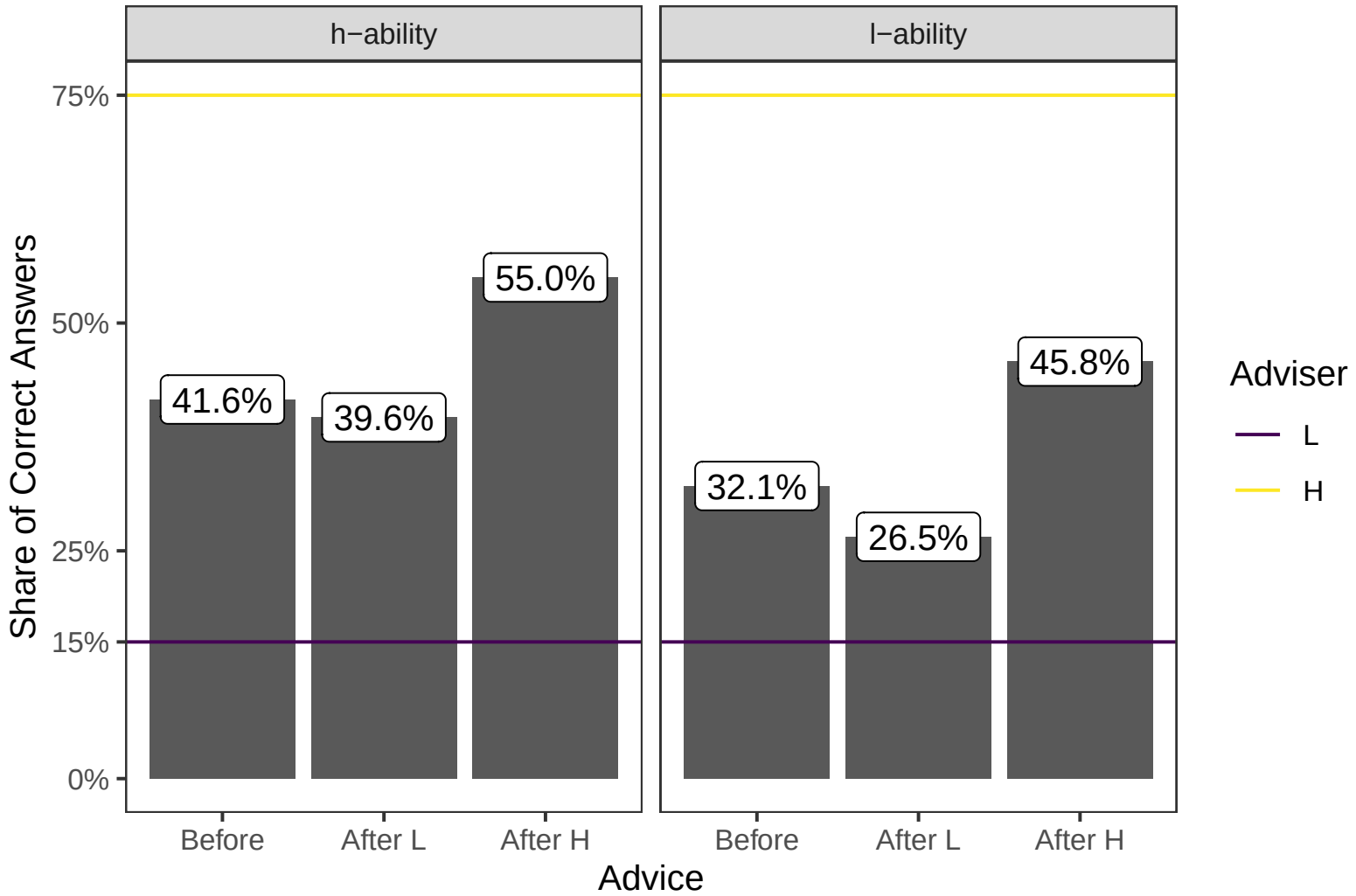
Elo


50%
40%
30%
20%
10%
0%
Share of Subjects
Unrated
(1000, 1100)
[1100, 1200)
[1200, 1300)
[1300, 1400)
[1400, 1500)
[1500, 1600)
[1600, 1700)
[1700, 1800)
[1800, 1900)
>=2000
Rapid Elo

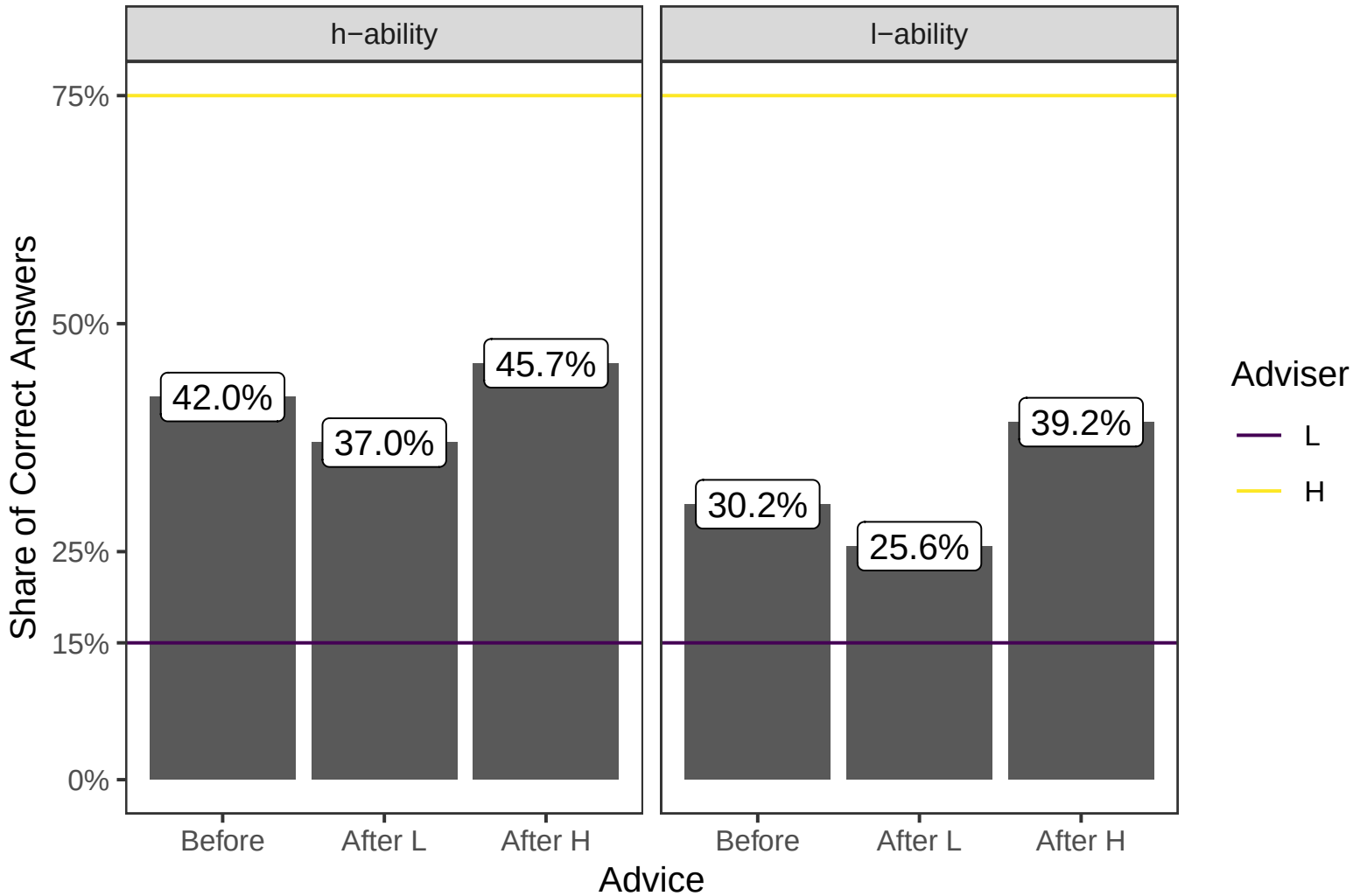
Back

Knowing the Advice Quality

Back



Ignoring the Adviser Quality



Back

Subject	Adviser	Before	After	P-value[1]
Unknown Adviser				
l	L	28.8%	25.6%	0.422
l	H	31.6%	39.2%	0.076
h	L	43.9%	37.0%	0.129
h	H	40.0%	45.7%	0.221
Known Adviser				
l	L	30.0%	26.5%	0.382
l	H	34.2%	45.8%	0.007
h	L	41.1%	39.6%	0.731
h	H	42.1%	55.0%	0.002

[1] P-value of the two-sided t-test of equal share of correct answers before and after advice.

Table: Share of correct answers, depending on the treatment, the subject and advisers type.

Back

Subjects Keep Their Answers

Subject	Low-quality advice	High-quality advice	Unknown advice
Lower-ability	80.4%	61.5%	69.4%
Higher-ability	85.0%	69.6%	83.5%

Table: Proportion of kept answer, by treatment

When Agree Back

Subjects	Treatment	Agree[1]	When Agree Before[2]	
			Keep	React
Lower-ability	Know H	33.5	93.0	7.0
	Know L	27.8	98.6	1.4
	Unknown	29.4	94.0	6.0
Higher-ability	Know H	41.8	96.5	3.5
	Know L	26.8	95.9	4.1
	Unknown	33.0	97.3	2.7

We remove from this table the missing answers because we have no distance from the answer for them. We therefore slightly overestimate the agreement percentage before receiving the advice.

[1] Percentage of identical pre-advice answers with the adviser.

[2] Percentage of kept or changed answer conditional on pre-advice answer being identical to the adviser's.

Table: What happens when subjects agree with advice.

Back

Heuristics

- First-best: After receiving advice, you know what is the right answer and can choose either your previous answer or the adviser's
- Probabilistic: You know your own and the adviser's probability of success in each round and choose the answers with the highest probability
- Elo: You follow the advisers if their Elo is higher than yours

Heuristics

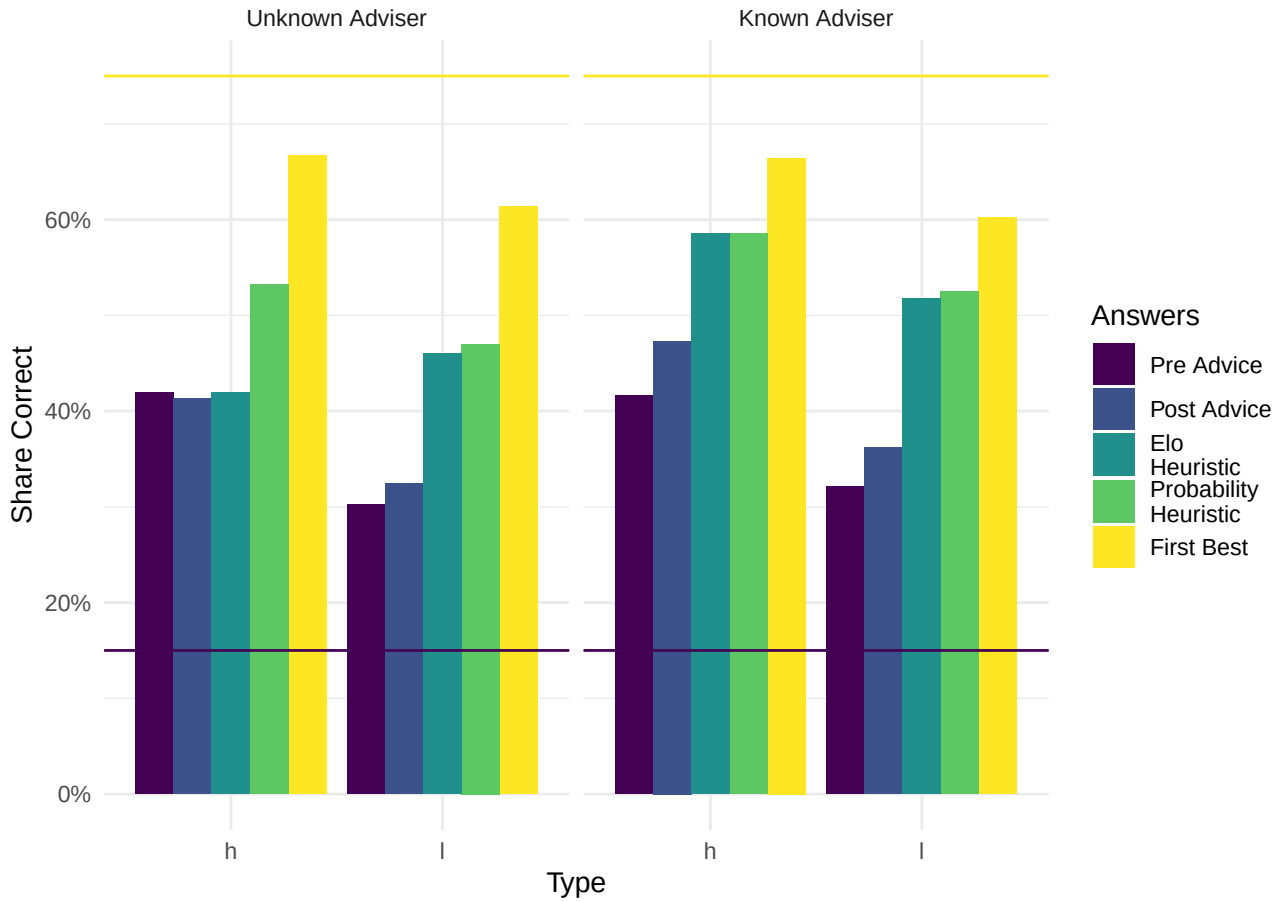


Figure: Share of correct answers following different heuristics.

Back

Elo Heuristic

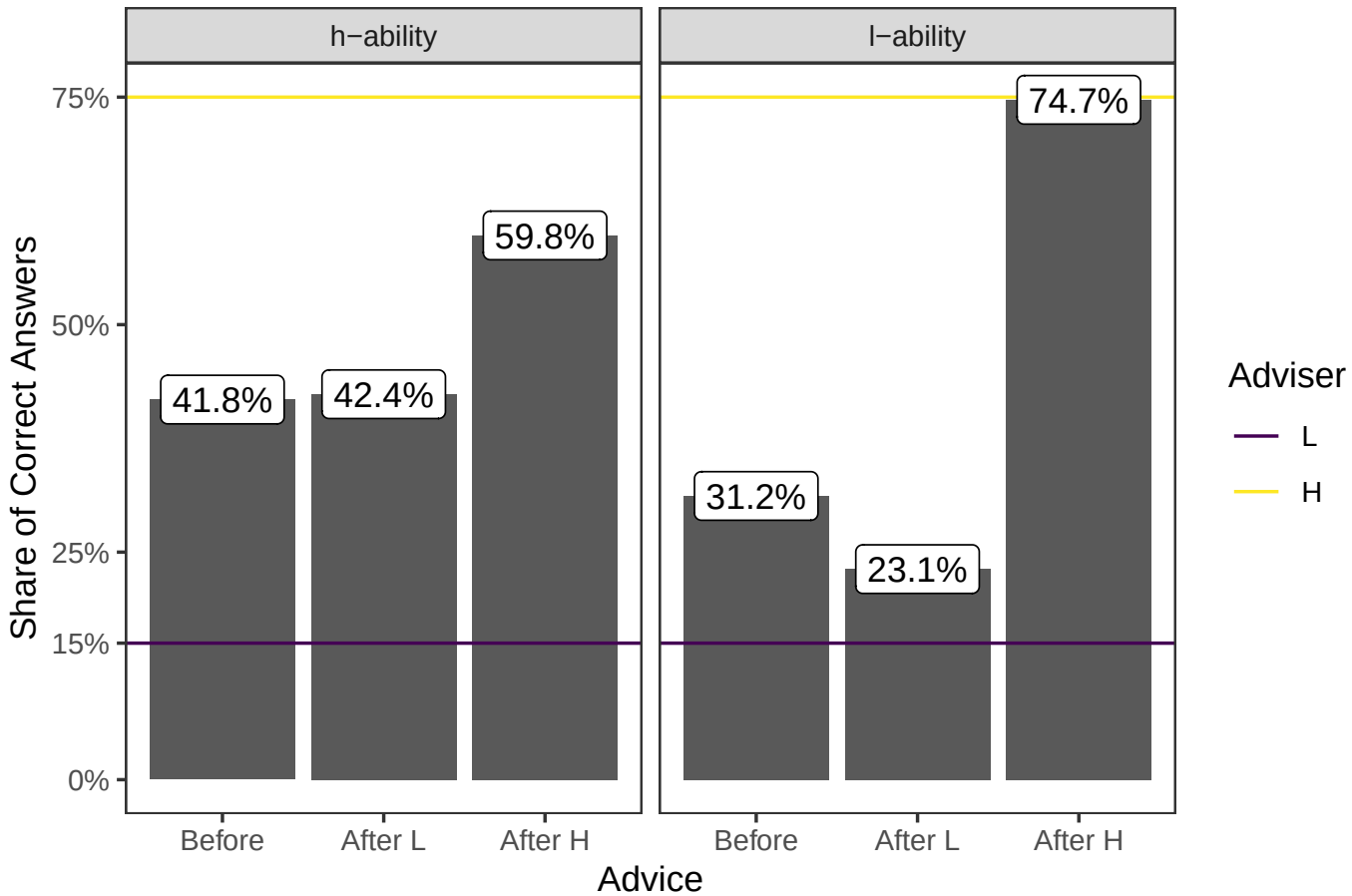


Figure: Share of correct answers following the Elo heuristic.

Back

"First-best"

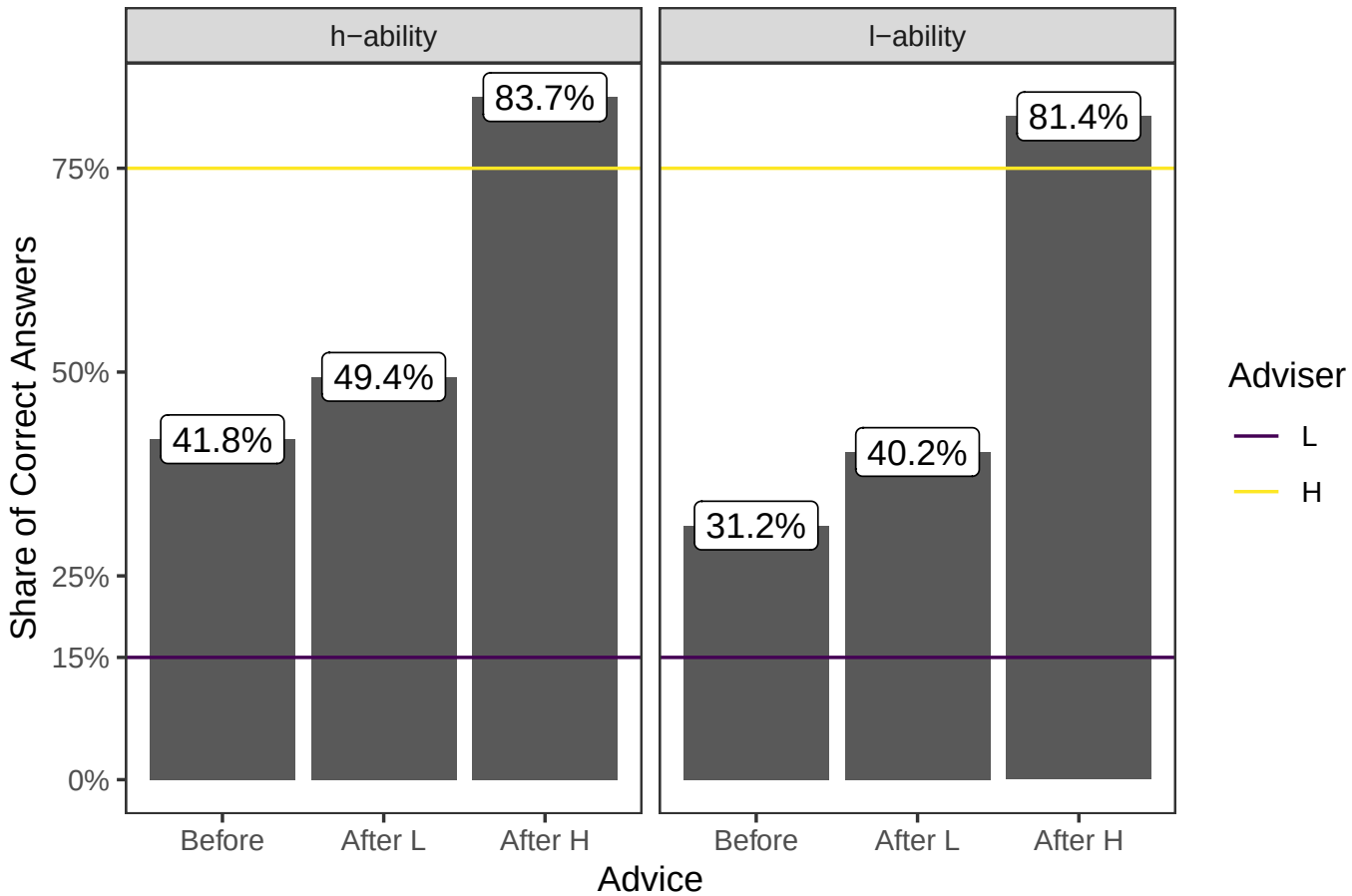


Figure: Share of correct answers in the first-best case.

Back