Decreasing Differences in Expert Advice: Evidence from Chess Players

Elias Bouacida, Renaud Foucart, Maya Jalloul

Decreasing Differences in Expert Advice: Evidence from Chess Players *

Elias Bouacida[†] Renaud Foucart[‡] Maya Jalloul[§]

February 2, 2024

Abstract

We study the impact of external advice on the relative performance of chess players. We asked players in chess tournaments to evaluate positions in past games and allowed them to revise their evaluation following advice from a high or a low ability player. While our data confirms the theoretical prediction that high-quality advice has the potential to act as a "great equalizer," reducing the difference between high and low ability players, this is not what happens in practice. This is in part because our subjects ignore too much of the advice they receive, but also because low ability players pay – either due to overconfidence or intrinsic preference – a higher premium than high ability ones by following their initial idea instead of high-quality advice.

Keywords: decreasing differences, expert, advice, chess, control
JEL-Codes: C78, C91, C93, D91, J24, O33

*We thank Jamal El Chamieh and Omar El Deeb for their support in organizing the experiment, Alexander Matros and David Rietzke for taking the time to solve our positions, and participants to seminars and workshop in Lancaster and at the 1st Doctoriales ASFEE.

[†]Université Paris 8, elias.bouacida@univ-paris8.fr
[‡]Lancaster University Management School, r.foucart@lancaster.ac.uk
[§]Lancaster University Management School, m.jalloul@lancaster.ac.uk

1 Introduction

Does offering high quality advice help reduce the gap between good and bad experts? In theory, the answer is yes: the benefit from being able to rely on outside advice is higher if your own expertise is lower. In a pre-registered[1] lab-in-the-field experiment with chess players participating in tournaments in Lebanon, we find that it is only true to a limited extent. High quality advice has the potential to benefit low ability players significantly more than high ability ones, but in practice it fails to make any significant difference. Our subjects reveal such a high preference for following their first idea and ignoring additional information that they forego a large share of the potential gains from the advice. Lower ability players end up paying the highest premium from ignoring good advice.

A major promise of the increasing digitalization of sectors such as legal studies, computer science, cancer diagnostic, or surgery, is to allow average practitioners to benefit from increased access to the knowledge of the very best in their fields. This is particularly true with the emergence of Artificial Intelligence (AI), which has shown its potential to improve the performance of 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) in routine tasks.

Perhaps the most striking result from this emerging literature is the potential for such advice to act as a "great equalizer:" if everyone has access to the same high quality external input, this should indeed benefit us all, but mostly those with lower expertise in the first place. This property is known as *decreasing differences* and theoretically holds for most ways of matching expertise (Chade and Eeckhout, 2018): the marginal impact of the quality of advice is decreasing in the ability of the expert who receives it. Other examples include production in garment factories (Hamilton *et al.*, 2003; Adhvaryu *et al.*, 2020) and student coursework in universities (Fischer *et al.*, 2023). A more general study of the US labour market shows that lower ability workers typically benefit more from the being part of a team with a high ability partner (Herkenhoff *et al.*, 2024).

The decision to ignore one's own signal and follow the advice of others is typically studied in economics in the context of information cascades (Anderson and Holt, 1997; Kübler and Weizsäcker, 2004), and there is evidence that subjects often like to bet on themselves even when it is optimal not to do so (Weizsäcker, 2010). In psychology, a large literature studies how subjects tend to give a sub optimal weight on advice in their decision-making (Bailey *et al.*, 2022; Bonaccio and Dalal, 2006). This result is also linked to the idea of preference for decision rights or control premium (Bartling *et al.*, 2014; Owens *et al.*, 2014) ; and the "illusion of control" (Langer, 1975; Sloof and von Siemens,

[1]The pre-registration is available at: https://aspredicted.org/124_MSY

2017) where subjects are overconfident when they make the decision themselves. A recent literature also studies whether people react differently to human advice and to the one provided by AI (Candrian and Scherer, 2022; Hertz and Wiese, 2019).

We partnered with a local academy to run our experiment alongside chess tournaments in several locations in Lebanon in the Summer of 2023. We paid the participation fee of our subjects to the tournament, as well as variable monetary incentives depending on their performance. The main task our subjects had to perform was to evaluate the *pawn advantage* of 20 chess positions – a measure of which player is better positioned to win the game, and by how much – taken from a large database of past games. For each position, we first asked our subjects to make their own evaluation, by choosing one of four possible answers. We then provided them with the evaluation of an external adviser for the same position, and asked our subjects to evaluate it again. Subject choices were not visible to the experimenters and remained anonymous. Our paper is thus closer to the literature on advice than control, as our subjects could not delegate their decision to the adviser, but we offered them to reconsider their own decision based on external advice. Hence, they had to write their evaluation in the same way regardless of whether they took the advice into account. Both pre- and post-advice evaluations were selected for payment with equal probability.

One of our advisers is an International Master, with an Elo rating placing him among the top 6,000 players in the world, and better rated than all of our subjects. The second adviser is an everyday chess player with no formal rating, placing him at the bottom of our subjects. In one treatment, we disclosed the rating of both advisers, but only told our subjects the advice came from one of them with equal probability. In the other treatment, we also informed the subjects of which expert the advice came from.

As expected, the advice of our International Master (75% of correct answers) is much better than the advice of our everyday player (15% of correct answers). As from our pre-registration, we defined as "high ability" our subjects with an official FIDE rating in the top half of our sample, and as "low ability" those in the bottom half. The results are similar if we define as "high ability" those with a rating and low ability the unrated, as 55 out of our 102 subjects had an official FIDE rating. Before receiving the advice, high ability subjects had a rate of correct answers of 41.8%, and low ability ones of 31.2%.

In our main pre-registered test, we find no evidence of decreasing differences, with a slightly and non-significantly higher share of good answers when the matching of subjects with advisers is disassortative – low ability subjects with our best adviser and high ability ones with our worst one – than in the opposite case (40.5% versus 38.4%). We then ran a similar exercise by comparing the rate of good answers when observing our best adviser as compared to no advice at all, and do not find any statistically significant difference

either (42.5% versus 40.1%).

The potential for decreasing differences in our setting is however real, and strikingly different from the observed behaviour of our subjects: by blindly following all advice, no subject would answer correctly less than 45% of the time. By following exclusively the advice of the International Master and answering at random when the adviser is the everyday player, one could expect 50% of correct answers. Yet, our subjects display a strong preference for following their own idea: a large majority of the evaluations remain unchanged after observing advice different from their initial evaluation. Even among our lower ability subjects, 52.9% of the evaluations remain unchanged after observing a different advice from an International chess Master.

Our paper contributes to the scientific literature on decreasing differences, advice and control discussed above. The main novelty of our research is to explicitly measure the potential for decreasing differences in a context where our subjects have a certain level of expertise on their topic and have to confront that expertise with a possibly better external advice. It serves as a cautionary tale for views that greater access to advice will compress the productivity distribution and reduce inequality between workers: while good advice may benefit the less able more, there is no reason to believe that the ability to identify and follow good advice is homogeneously distributed in the population. If the most able are also the most able to use advice, it is far from obvious that inequality is reduced.

We also contribute to the literature on control and advice by providing results from a non-WEIRD (White, Educated, Industrialized, Rich, and Democratic) sample (Henrich *et al.*, 2010), as our subjects live in a Middle-Eastern country in the midst of a banking and political crisis. Finally, this paper is part of a literature using chess players to study human decisions, such as strategic behaviour in sequential games (Levitt *et al.*, 2011), gender differences in risk-taking (Gerdes and Gränsmark, 2010), social norms and the gender gap (Dilmaghani, 2021), or the role of superstars (Bilen and Matros, 2023).

The rest of the paper is organised as follows. We start in Section 2 by providing some background about the nature of the tasks, our measure of ability, and why we should expect to observe decreasing differences in the impact of the quality of advice. In Section 3, we describe our experimental protocol, procedures and pre-registered outcomes. We present the results in Section 4 and conclude in Section 5.

2 Tasks, Ability, and Decreasing Differences

In this section, we describe our main task of evaluating a position in terms of *pawn advantage*. We then explain the measure we use to rank subjects by ability group, the *Elo*

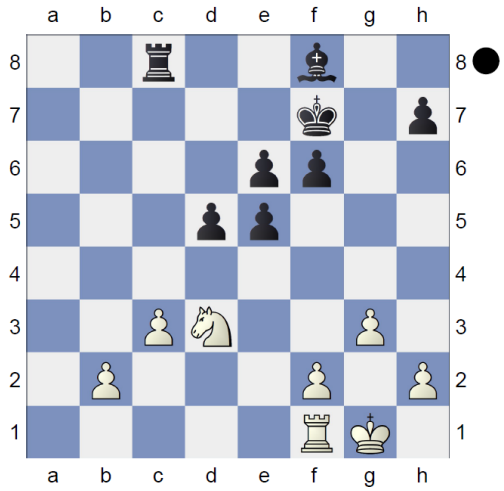


Figure 1: A position in a game of chess, as shown to our subjects.

rating. Finally, we outline the theoretical argument for why we should expect decreasing differences in our context.

2.1 Pawn Advantage

We asked subjects to evaluate chess positions, which is a description at a given point of a game of the positions of the pieces on the board. We show in Figure 1 one of the positions used in our experiment, exactly as we showed it to our subjects. Positions are evaluated using the notion of *pawn advantage*, a measure of which player (White or Black) is better placed to win the game. We chose 20 positions from past games of chess using the Chessbase Mega database 2023, and picked half of them with a pawn advantage of 0.7 (a slight advantage) and the other half with 2.4 (a large advantage), either for Black (-2.4 and -0.7) or for White (0.7 and 2.4).[2] The task for our subjects was to identify the correct evaluation out of the four possible ones (in Figure 1, the correct answer is -0.7). While evaluations are not an exact science, contemporary chess engines converge towards almost identical pawn advantages. There is thus no ambiguity as to which of the four evaluations is the correct answer.

[2]Chess players are in general reluctant to translate pawn advantages into winning probabilities, one reason being that there are not two but three possible outcomes in chess: a win, a loss, or a draw. According to one measure however (suggested by Sune Fischer and Radu Pannan based on 405,460 past games), a pawn advantage of 0.7 correspond to a 60% probability of win and of 2.4 to a 80% probability of win - counting a draw as half a win. In our selection of positions, we followed this statistical regularity: of the games with a pawn advantage of ±2.4, 7 ended with a win for the advantaged player, 2 with a draw, and 1 with a loss ; of the games with a pawn advantage of ±0.7, 3 ended with a win for the advantaged player, 6 with a draw, and 1 with a loss.

To measure the influence of advice on evaluations, subjects were asked to evaluate positions first without advice, and then got a chance to revise their answer after seeing the advice of either an unrated player or of a highly rated player, based on a measure called the Elo rating.

2.2 Elo Rating

In order to rank our subjects and advisers by their estimated ability, we use their *Elo Rating*, a system created by Arpad Elo to compute the relative skill level of a player. When two players play against each other in a tournament registered with the international chess federation FIDE, the winner gains Elo points, and the loser loses points. The number of points gained and lost depends on the difference in ratings and on the expected outcome. Any player with a rating strictly lower than 1000 is considered as unrated by the FIDE (and in our sample). As a rule of thumb, a difference of 100 points in the Elo rating means that the best rated player is expected to win 5 out of 8 games. While Elo is an imperfect measure of ability, it is taken seriously by players. In our case, the average rating of our rated players is 1490, while our best subject is in the range 2100-2200. We plot the Elo distribution of our subjects in Figure 5 in Appendix A. It should thus be clear to all our subject that our high ability adviser with a rating higher than 2300 is more likely than them to correctly evaluate a position. It should also be clear that our low ability adviser, described as an "an unrated player, who plays regularly for fun" is not of strictly higher ability than any of our subjects who all take part in a registered tournament.

2.3 Decreasing Differences in the Return from Expert Advice

Consider two subjects l and h, with perfect information about their own probability of successfully solving a task p_i, $i \in \{l, h\}$, as well as the probability of the low L and high H ability advisers to do so $q_L < q_H$. In the case in which subjects do not know the identity of the adviser – but know both are equiprobable – we denote by $\bar{q} = \frac{q_L+q_H}{2}$ this probability.

Unless all subjects follow (or ignore) all types of advice, we should observe strictly *decreasing differences* if subjects correctly infer the probabilities and maximize their expected probability of finding the correct answer.

Define by $f(i, j)$ the probability that subject i solves a task correctly after observing advice j and assume that $q_L < p_l < p_h < q_H$. If subjects want to maximize their probability of success and know the identity of the adviser, $f(l, L) = p_l$, $f(h, L) = p_h$, and $f(l, H) = f(h, H) = q_h$.

It is easy to see that in that case, the function displays decreasing differences:

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

as the expression simplifies to $p_l < p_h$. This statement is equivalent to saying that Negative Assortative Matching (NAM) of subjects to advisers yields a higher expected share of correct answers than Positive Assortative Matching (PAM),

$$\frac{f(l,H) + f(h,L)}{2} > \frac{f(l,L) + f(h,H)}{2}.$$

The same result holds when considering the case of unknown advisers if $p_l < \bar{q} < p_h$, so that type l subjects follow all advice and type h do not follow any. In that case, $f(l,L) = q_L$, $f(h,L) = f(h,H) = p_h$, and $f(l,H) = q_H$. The condition for decreasing differences is then $q_H > q_L$, and the difference between NAM and PAM is higher than with known advisers. The reason is that a good adviser then not only helps more the low ability subjects, but it also protects them from following bad advice. Finally, if $\bar{q} \geq p_h$ or $\bar{q} \leq p_l$, the differences are constant and the probability of a correct answer in NAM is the same as in PAM. This result is trivial, as it simply states that if all subjects follow all advice, they also solve all problems with the same probability, and if they ignore all advice, the quality of advice has no influence on their success.

By the same logic, we can compare advice from H and no advice at all, where $f(i,0) = p_i$ is the probability of the answer of subject i being correct before advice. With known adviser, the result is identical to the one above, as $f(i,L) = p_i$ for both types of subjects. With unknown adviser, there are always decreasing differences unless all advice is ignored. If $\bar{q} < p_h$, the condition becomes $q_H > p_l$. If $\bar{q} \geq p_h$, it is $p_h > p_l$.

There are however two main biases and preferences that could influence our theoretical result of decreasing differences in the experiment. The first is that our subjects do not have full information on their probability of success and the one of their advisers. If lower performing subjects are also more overconfident than high ability ones, they may benefit relatively less from advice. The second is preference for following their initial idea: if lower ability subjects value more strongly keeping their first answer than high ability ones, they are less likely to follow advice for a given expected gain, decreasing the potential for advice to act as a great equalizer.

3 The Experiment

We ran the experiment during the Summer of 2023 in several cities in Lebanon, alongside tournaments organised by a local academy.[3] Our subjects were regular participants in

[3]The exact dates are August 15, August 20, September 2, and September 17, 2023. In line with the pre-registration, we stopped recruiting participants when we reached 100 subjects, so that we recruited

tournaments, and had therefore a certain level of expertise in the game. We describe their self-reported demographic characteristics in Table 6 in Appendix A. All subjects received the experimental material written both in English and Arabic.

3.1 Protocol

We recruited subjects before the tournament through the organizing chess academy and paid for their registration (around $5) as a participation fee. The experiment took part in a separate room at times where our subjects were not playing. Each subject was randomly allocated either to a treatment with or without information on the adviser. Subjects received tasks booklets and answer sheets (see Figure 2) upon being seated. There were two rounds of tasks, each corresponding to solving ten positions. In each round, subjects were given 8 minutes to complete their evaluation of each position among the possible choices ($-2.4, -0.7, +0.7$, or $+2.4$) on the left part of their answer sheets. Then, they were provided with the answers of one of our two advisers for the same questions (Figure 3). They were given 4 minutes to look back at their answers, compare with the advice, and complete the right part of the answer sheet with their possibly updated evaluations.

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								

Figure 2: The answer sheet for the first round. The left-hand side was completed before seeing the advice, and the right-hand side after.

The tasks booklets were labelled 1 or 2 corresponding to the known or unknown adviser condition, and a or b corresponding to the order in which the advice was received, where a means the H adviser for the first round of ten evaluations and the L adviser for the second

a total of 103 subjects. Our total sample is however $n = 102$ as, in line with our pre-registration, we removed observations for which no choice were made and one of our subjects did not write anything in the second part of the answer sheet. The project has received IRB approval from Lancaster University.

one. In the known adviser condition, we told subjects that the answers we gave them were coming from "a player with a rating of 2335" for the round in which their adviser was H, and from "an unrated player, who plays regularly for fun" when their adviser was L. In the unknown adviser condition, we told them that "With equal probability, the player has a rating of 2335, or it is an unrated player, who plays regularly for fun".

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.

Figure 3: A sheet containing advice for one round.

After solving the two rounds of evaluations, subjects completed a short demographic questionnaire as well as questions about their stated preference for control (5 questions borrowed from Burger and Cooper, 1979). We provide all the experimental material in Appendix C. All the sessions were administered by one of the co-authors of this study (Maya Jalloul), who read the experimental material and ensured no one could cheat.

On top of the participation fee, we picked one of the 40 evaluations of each subject at random (20 evaluations before advice, and 20 after) and paid a variable amount of \$10 if the answer was correct.[4] We only knew the subject number, and not their identity. We communicated a list of payments and subject numbers to the organizing chess academy, who then processed the payments based on a list they made allocating participant numbers to individuals. At the end of the experiment, they destroyed these lists.

3.2 Pre-Registered Hypotheses

Our main pre-registered outcome was the existence of decreasing differences, tested by comparing the probability that an answer is correct under Negative Assortative Matching

[4]Given the difficult banking situation in Lebanon and the fact that some of our subjects were minor, we did not pay subjects directly in cash but with monetary vouchers for subsequent tournaments or other spending on the day.

of subjects to advisers with the probability that it is correct under Positive Assortative Matching.

In line with our pre-registration, we divided our sample of $n = 102$ into two groups of equal size, based on their Elo rating, and removed those questions for which subjects did not answer. As 54 subjects had a formal Elo rating, the results are almost identical when considering a dichotomy rated/unrated instead. The subjects in the lower ability group are denoted as l and those in the higher ability group as h. We verify empirically that our division reflects an average difference in ability to solve our tasks by comparing their respective share of correct answers pre-advice (see Table 1). We also see that, on average, our L adviser is less likely to answer correctly than our l subjects, who are themselves less likely to do so than h subjects. Finally, our H adviser performs much better than the average h subject, so that on average $q_L < p_l < p_h < q_H$.

Table 1: Share of correct answers before advice.

Player type	
Subjects l	31.2%
Subjects h	41.8%
Adviser L	15.0%
Adviser H	75.0%

We averaged the share of correct answers of l subjects after observing H advice and of h subjects after observing L advice to have our measure of Negative Assortative Matching. We then did the same for PAM with l having observed L advice and h having observed H advice. Our main test was to see if, as in Section 2.3, NAM indeed generates a significantly higher share of correct answers than PAM. Our alternative main pre-registered test was to do the same exercise using the pre-advice evaluation instead of the revised evaluation following L advice. We had also pre-registered to look at the same question for the two experimental conditions separately.

4 Results

4.1 Individual Performance

We present the result of our main pre-registered test in the first part of Table 2. Looking at all subjects accross treatments, we see that matching high ability subjects with low ability advice and low ability subjects with high ability advice (NAM) yields on average 40.5% of correct answers, only slightly more than the 38.4% of correct answers when

matching subjects and advisers assortitatively (PAM). When looking at the two experimental conditions separately, we see that the difference is never significant, and that, when the adviser is known, the proportions are almost equal. In the second part of Table 2, we look at our alternative pre-registered test in which we replace low ability advice by the pre-advice evaluations, we see that the difference between NAM and PAM is not significant either.

While we cannot rule out that some decreasing differences exist, our sample size should have been sufficient to identify any large effect. To give an idea of our statistical power, in order to detect a significantly different proportion between PAM (35.21%) and NAM (38.12%) in the Unknown Adviser treatment, which has the largest effect, we would have needed a sample size of 4,284, whereas we aimed for a sample size of 1,000 and our realized sample size is 960 in that treatment. The other way to look at it to see what effect size we can measure with the sample size we have. With the proportion of PAM we currently have, we would be able to detect a significant ($p < 0.05$) effect for a proportion with NAM at 41.42%.

Following the logic of Section 2.3, we can get an intuition of the lower and upper bound for the difference between NAM and PAM if payoff-maximizing subject followed a simple decision rule. As an upper bound, in the unknown adviser treatment, if all l-subjects follow all advice, and all h-subject ignore all advice, the difference between NAM and PAM would be of 30 percentage points. As a lower bound, in the known adviser treatment, if all subjects ignore L-advice and follow H-advice, the difference would be of 5 percentage points. This contrasts with the non-statistically significant difference of 2.5pp and 2.7pp found in our experiment.

Pooling across treatments, both our high ability (from 41.2% to 50.8%) and low ability (from 32.9% to 42.5%) subjects see their share of correct answers increase by 9.6 percentage points after seeing good advice. After seeing advice from our low ability adviser, the share of good answers of our low ability subjects drops by 3.3pp, to 26.1%, while for our high ability subjects the drop is by 3.9pp to 38.4%. The negative impact of L-advice is largely a mechanical result following from the fact that our low ability adviser performed worse than the expected result of someone answering at random.

To see why despite a large initial difference in the ability of our subjects high quality advice fails to play the role of a great equalizer, we look at the share of evaluations remaining unchanged after observing advice. Perhaps the most striking fact is that advice remains largely ignored. A first descriptive statistics is that, overall, around three-quarters of the choices are unchanged after observing advice (we report in Table 7 in Appendix B.1 the figure for each treatment and type).

Unchanged advice can however be for several reasons, one of them being that if a

Table 2: Main pre-registered test: comparing the share of correct answers post-advice under Negative Assortative Matching (NAM) and Positive Assortative Matching (PAM) of subjects to advisers.

Treatment	NAM	PAM	P-value[1]
Main test: H vs L-advice			
All	40.5%	38.4%	0.365
Unknown Adviser	38.1%	35.2%	0.384
Known Adviser	42.6%	41.3%	0.711
Alternative test: H vs No-advice			
All	42.5%	40.1%	0.301
Unknown Adviser	41.5%	36.9%	0.165
Known Adviser	43.3%	43.0%	0.951

[1] P-value of the two-sided two sample test of equal proportion between NAM and PAM.

subject's answer is identical to the adviser's there is no reason to modify it. The first three columns of Table 3 indeed show, unsurprisingly, that when their initial answer is the same as the adviser, our subjects tend to keep it.[5] The last four columns show what happens when they differ. Participants mostly, and correctly, ignore advice from our L-adviser, although almost 10% of our high ability subjects update their evaluation following advice from an adviser they should expect to be worse than them, in line with Schultze *et al.* (2017) who shows that some subjects feel the need to incorporate even useless advice. Low ability subjects ignore advice from unknown advisers around two-thirds of the time, as compared to around 80% for the high ability subjects. Our low ability subjects ignore (46.8%) or even move further away (1.8%) from our H-adviser roughly half of the time, only slightly less than our high ability subjects. Given the large difference in the pre-advice results of both types of subjects, this choice either implies that low ability subjects are much more overconfident than high ability ones, or that they have a much larger intrinsic preference for not following advice.

To see how much payment subjects left on the table by ignoring advice, we compare in the next section their choices in our experiment with two "heuristics" of always following or ignoring some type of advice.

[5]We remove a small number of null answers from this table, because we have no distance from the answer for them. We therefore slightly overestimate the agreement percentage before receiving the advice.

Table 3: How do subjects react to the advice received, depending on agreeing or not with it?

		Agree	Agree Before[2]		Disagree	Disagree Before[2]			
Type	Treatment	Before[1]	Keep	React	Before[1]	Keep	Follow	Closer	Further
l	Know H	**33.5**	93.0	7.0	**66.5**	46.8	39.2	12.3	1.8
l	Know L	**27.8**	98.6	1.4	**72.2**	78.8	14.0	7.3	0.0
l	Unknown	**29.4**	94.0	6.0	**70.6**	64.3	27.3	6.3	2.2
h	Know H	**41.8**	96.5	3.5	**58.2**	52.5	42.5	2.5	2.5
h	Know L	**26.8**	95.9	4.1	**73.2**	83.9	9.5	4.0	2.5
h	Unknown	**33.0**	97.3	2.7	**67.0**	79.7	15.3	3.3	1.7

[1] Bold font: percentage of identical and different pre-advice answer than the adviser.

[2] Normal font: percentage of kept or changed answer, conditional on pre-advice answer being identical to adviser's; percentage of kept or changed answer (following, getting closer, or further away from advice) conditional on pre-advice answer being different from adviser's.

4.2 Heuristics and Premium for Ignoring Advice

Our first "Elo" heuristic is arbitrary and purely deterministic. We assume that, when they know the adviser type, all subjects follow H advice and ignore L advice. When the adviser is unknown, our high ability subjects ignore all advice, and our low ability subjects follow all advice. While this approach has the advantage of being simple and corresponds to information known ex-ante by the subjects, it does not rely on actual probabilities of success.

Our second "Probability" heuristic is based on actual probabilities, an information not available to our subjects. We approximate a subject i's probability of evaluating a position correctly p_i by their share of correct answers pre-advice, and, similarly, the probability for experts to do so q_L and q_H, with $\bar{q} = \frac{q_L+q_H}{2}$ the probability for unknown advice. This approach differs from the one of section 2.3, as we do not look at different probabilities at the question level, but at the individual level for a given round of ten positions, so that this "first-best" way of incorporating advice follows a simple decision rule: if $p_i > q_H$, ignore all advice ; if $p_i \in (\bar{q}, q_H)$, only follow the known advice of H ; if $p_i \in (q_L, \bar{q})$, follow all advice, known or unknown, except for the advice of L ; and if $p_i < q_L$, follow all advice.

For each subject, we pick these heuristics and see how many correct answers they could have achieved by following them. While imperfect (and not part of our pre-registration),

Table 4: Difference (in percentage points) between the average share of correct answers of h and l type subjects having received H-advice, following our heuristics and in the experiment.

		Types		
Treatment	Heuristics	l	h	P-value[1]
Unknown	Probabilistic	32.4	23.0	0.056
Known	Probabilistic	27.7	21.1	0.134
All	Probabilistic	30.0	22.2	0.014
Unknown	Elo	36.8	-5.7	<0.001
Known	Elo	27.7	21.1	0.139
All	Elo	32.2	9.0	<0.001

Premia paid for ignoring H-advice are given in percentage points.

[1] P-value of the two-sided two sample t-test of equal control premium between the h and l Elo types.

this method gives us an illustrative idea of the potential of advice and its role as a great equalizer. We cannot rule out that subjects could have done even better if they were able to identify the questions for which they are particularly confident to have a correct answer for instance.

We measure in Table 4 the premium subjects are paying in order to ignore high-quality advice, defined as the difference (in percentage points) in the share of correct answers post H-advice if they followed our heuristics and in the experiment. Accross treatments, our low ability subjects would have a 30 percentage points higher share of correct answers after H-advice following the Probabilistic heuristics than they did in the experiment, as compared to 22.0 for our high ability subjects. Using the Elo heuristic, these number are 32.2pp and 9.0pp respectively. Both differences are statistically significant. We can thus conclude that, be it because of overconfidence or intrinsic preference for keeping their original answer, our low-ability subjects ended up paying a higher premium than high-ability ones for ignoring good advice.

We then aggregate both types of advice to see how this result is affected by the possibility of receiving low quality advice. We plot on Figure 4 the share of correct answers before advice, post-advice in the experiment, and post-advice following our two heuristics. The left-hand side corresponds to the treatment with unknown adviser. There,

our two heuristics are radically different. The reason is that our H-adviser performed so much better than all our subjects that many of our best subjects would have been better off following all advice. This implies that, following our Elo heuristics, advice is more than a great equalizer, it leads to our low-ability subjects performing better than the high-ability ones. Under our probabilistic heuristics however, our best subjects continue to perform better than the low ability ones, even after following advice. This is because, in order to benefit from good advice, subjects have to also follow some very bad one – as our L-adviser performed worst than randomness – and decrease their performance. Thus, the top performers among our high ability subjects ignore advice and continue to overperform.

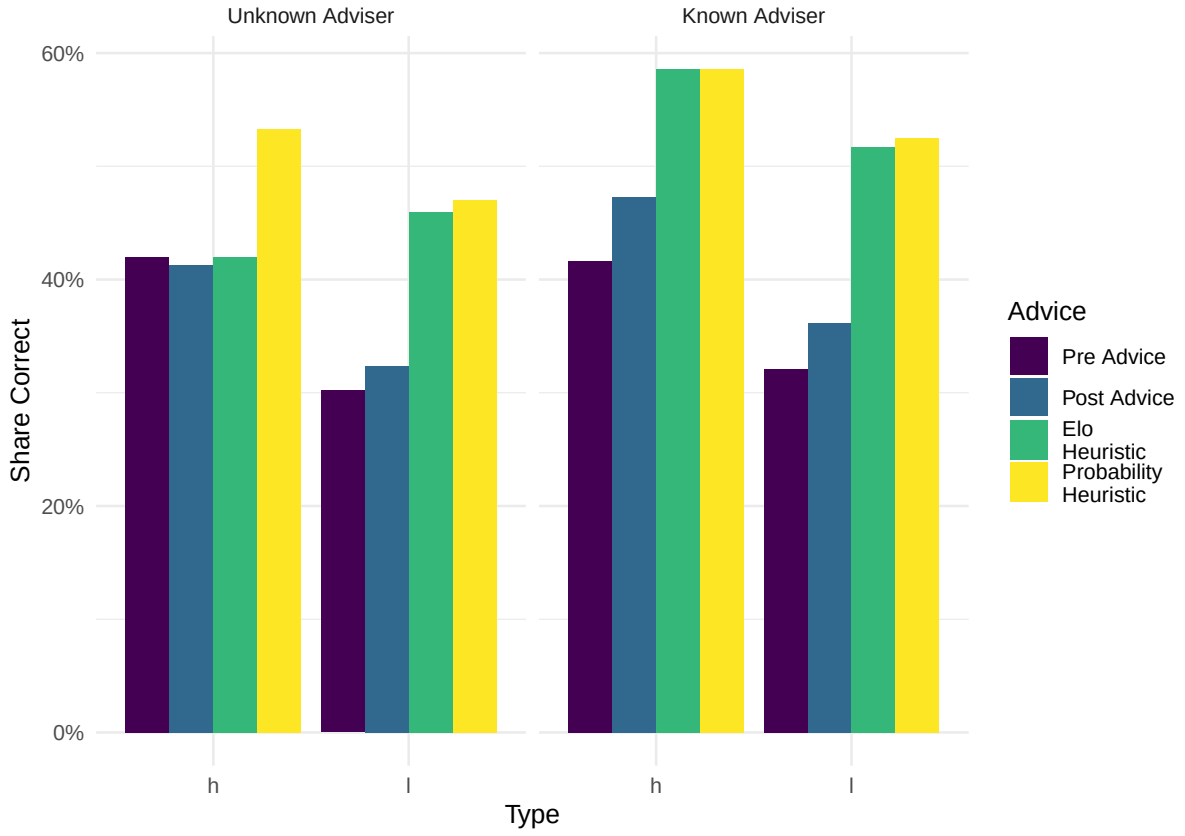


Figure 4: Share of correct answers before advice, post advice, and following our heuristics, in our two experimental conditions

In the case with known adviser, the two heuristics are roughly similar: it is in the interest of almost everyone to ignore L-advice and to follow H-advice. The initial difference between the share of correct answers of high and low ability subjects is of 10 percentage points, and increases to 10.9pp post-advice in the experiment, as low ability subjects do not follow enough H-advice while nonetheless following some L-advice. In contrast, by following our heuristics, this difference could have decreased to between

5.8pp (Probabilistic heuristic) and 6.6pp (Elo heuristic). We report those differences for both treatments in Table 9 in Appendix B.2.

4.3 Correct Answers and Preference for Control

We start by constructing an index of the stated preference for control by aggregating the answers to our questions borrowed from Burger and Cooper (1979). We find that stated preference for control is correlated with the probability of a subject keeping their answers, after controlling for subject and answer characteristics (see Table 8 in Appendix B.1).

We then present our main regression on Table 5. Following our pre-registration, we use both our binary definition of low and high-ability subjects, and a continuous measure based on the Elo rating, and run regressions for the known and unknown adviser treatment.

As expected, a high ability adviser typically benefits subjects, and the better rated subjects are more likely to evaluate positions correctly. The interaction term between the adviser type and the rating of our subjects gives an alternative measure of the existence of decreasing differences. As in the main tests with two categories of subjects, it is not significant. We control for individual characteristics in Table 10 in Appendix B.2.

Table 5: Regression for the share of correct answers, with fixed effects at the position level.

	Grouped by ELO		Continous ELO	
	Known	Unknown	Known	Unknown
H Adviser	0.147 (0.021)	0.066 (0.192)	0.23 (0.002)	0.17 (0.010)
	(0.058)	(0.049)	(0.065)	(0.059)
Low Elo	-0.140 (0.001)	-0.118 (0.053)		
	(0.036)	(0.057)		
Elo			1.1×10^{-4} (<0.001)	8.7×10^{-5} (0.027)
			(2.2×10^{-5})	(3.6×10^{-5})
H Adviser×Low Elo	0.056 (0.267)	0.096 (0.093)		
	(0.049)	(0.055)		
H Adviser×Elo			-6.8×10^{-5} (0.071)	-7.2×10^{-5} (0.056)
			(3.5×10^{-5})	(3.6×10^{-5})
Std.Errors	by: position	by: position	by: position	by: position
Num.Obs.	1064	912	1064	912
R2	0.145	0.082	0.147	0.085
R2 Adj.	0.127	0.059	0.129	0.063

Notes: Robust standard errors clustered at the position level. In parenthesis on the same line are the p-value, below the standard error.

5 Conclusions

Digitalization and the development of Artifical Intelligence promise to give broad access to high quality specialist advice. In theory, one of the main consequences of this evolution is a compression in the distribution of productivity, reducing the difference between the best and worst performers. However, the literature on advice and preference for control tells us that subjects may simply not take up this advice.

In this paper, we used a sample of subjects with specialist knowledge in their topic in a natural setting – chess players evaluating chess positions during a chess tournament – to learn more about the "great equalizer" potential of advice. While we find evidence that improving the quality of advice could benefit low ability players more, most of the potential benefit of advice is wasted by subjects choosing to keep their initial evaluation. This preference for following their own expertise hurts low ability subjects the most, as they had the most to gain. The fact that low-ability subjects are also those paying – either due to overconfidence or intrinsic preference – the highest premium to follow their initial evaluation is consistent with the idea that the most able subjects are also the most able to follow advice. It could also be the case that ability in chess is not exogenous, and that the best rated players are precisely those who were able to listen to advice during their training.

Our experiment suggests that existing evidence of decreasing differences based on the routine use of AI advice to improve writing or the gathering of information may not translate easily to sectors such as medicine or engineering where subjects have expertise and may want to trust it more at the moment of the decision than any advice they receive. Good advice however still plays the important role of protecting subjects from bad advice. In a world in which advice is, on average, correct, lower ability subjects are more at risk of using bad advice by accident. In terms of public policy, this means that digital literacy and the importance of telling good from bad advice will become even more important when people start routinely relying on it.

Among the limitations of our paper is the fact that we do not distinguish between advice from humans and from computers. We did so because, since the landmark victory of chess engine Deep Blue versus the then world champion Garry Kasparov in 1997, chess players see algorithmic analysis of the games as the gold standard. This is precisely the reason why we could use the chess engines evaluation of the pawn advantage in our chosen positions as the unambiguously correct answer. Further studies of subject specialists such as our chess players would benefit from comparing computer-based advice and human one and see whether decreasing differences are more pronounced with the latter.

6 Bibliography

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.

Anderson, L. R. and Holt, C. A. (1997). Information cascades in the laboratory. *The American economic review*, pp. 847–862.

Bailey, P. E., Leon, T., Ebner, N. C., Moustafa, A. A. and Weidemann, G. (2022). A meta-analysis of the weight of advice in decision-making. *Current Psychology*, pp. 1–26.

Bartling, B., Fehr, E. and Herz, H. (2014). The intrinsic value of decision rights. *Econometrica*, **82** (6), 2005–2039.

Bilen, E. and Matros, A. (2023). The queen's gambit: Explaining the superstar effect using evidence from chess. *Journal of Economic Behavior & Organization*, **215**, 307–324.

Bonaccio, S. and Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational behavior and human decision processes*, **101** (2), 127–151.

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

Burger, J. M. and Cooper, H. M. (1979). The desirability of control. *Motivation and emotion*, **3**, 381–393.

Candrian, C. and Scherer, A. (2022). Rise of the machines: Delegating decisions to autonomous ai. *Computers in Human Behavior*, **134**, 107308.

Chade, H. and Eeckhout, J. (2018). Matching information. *Theoretical Economics*, **13** (1), 377–414.

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

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).

Dilmaghani, M. (2021). The gender gap in competitive chess across countries: Commanding queens in command economies. *Journal of Comparative Economics*, **49** (2), 425–441.

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

Gerdes, C. and Gränsmark, P. (2010). Strategic behavior across gender: A comparison of female and male expert chess players. *Labour Economics*, **17** (5), 766–775.

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.

Henrich, J., Heine, S. J. and Norenzayan, A. (2010). Most people are not weird. *Nature*, **466** (7302), 29–29.

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

Hertz, N. and Wiese, E. (2019). Good advice is beyond all price, but what if it comes from a machine? *Journal of Experimental Psychology: Applied*, **25** (3), 386.

Kübler, D. and Weizsäcker, G. (2004). Limited depth of reasoning and failure of cascade formation in the laboratory. *The Review of Economic Studies*, **71** (2), 425–441.

Langer, E. J. (1975). The illusion of control. *Journal of personality and social psychology*, **32** (2), 311.

Levitt, S. D., List, J. A. and Sadoff, S. E. (2011). Checkmate: Exploring backward induction among chess players. *American Economic Review*, **101** (2), 975–990.

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

Owens, D., Grossman, Z. and Fackler, R. (2014). The control premium: A preference for payoff autonomy. *American Economic Journal: Microeconomics*, **6** (4), 138–161.

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*.

Schultze, T., Mojzisch, A. and Schulz-Hardt, S. (2017). On the inability to ignore useless advice. *Experimental Psychology*.

Sloof, R. and von Siemens, F. A. (2017). Illusion of control and the pursuit of authority. *Experimental Economics*, **20** (3), 556–573.

Weizsäcker, G. (2010). Do we follow others when we should? a simple test of rational expectations. *American Economic Review*, **100** (5), 2340–2360.

Appendix

A Sample Description

Table 6 shows that most of our subjects are young men. Figure 5 that most of our subjects are rated, but the mode is not being rated. The proportion of unrated players means that the low Elo group is almost all made of unrated players.

Table 6: Demographic characteristics

Gender	
Female	10
Male	78
Undeclared	15
Age	
<18	34
18-29	42
≥30	14
Undeclared	13

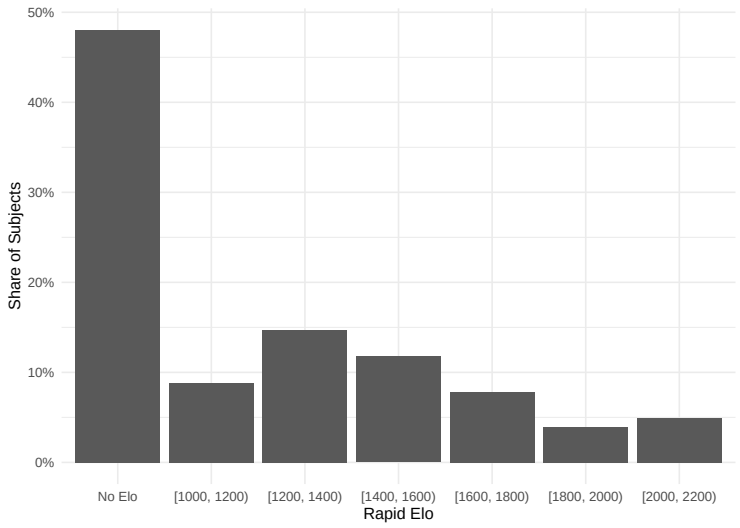


Figure 5: Distribution of Elo ratings among our subjects. The top rated adviser is above the upper limit.

B Additional Results

B.1 Keeping your Answer

Table 7 shows that subjects correctly change their answers more after seeing H-advice. They fail to change them as often as they should however. It also shows that lower Elo players change their answers more often that higher Elo ones, which is expected. Table 8 shows that higher Elo subjects tend to stick with their answers more often lower rated ones. When knowing the adviser, the baseline is that it is the L adviser, and subjects correctly keep their answer more often. On the other, as shown by the interaction terms between the H adviser and the Known Adviser, then subject change more often their answer when they know it is of good quality.

Table 7: Share of identical answers for l and h subjects after observing different types of advice.

Type	L-Advice	H-advice	Unknown Advice
l	80.4%	61.5%	69.4%
h	85.0%	69.6%	83.5%

Table 8: Regression for keeping the answer after receiving the advice, with fixed effects at the position level.

	(1)	(2)	(3)	(4)
Distance Correct[1]	-0.043 (0.001)	-0.0475 (0.001)	-4.0e-02 (0.002)	-4.5e-02 (0.002)
	(0.012)	(0.0123)	(1.1e-02)	(0.01212)
H Adviser	-0.019 (0.544)	-0.0085 (0.795)	-1.9e-02 (0.544)	-8.2e-03 (0.808)
	(0.031)	(0.0322)	(3.1e-02)	(0.03339)
Known Adviser	0.066 (0.009)	0.0414 (0.089)	7.9e-02 (0.004)	4.9e-02 (0.052)
	(0.022)	(0.0230)	(2.4e-02)	(0.02387)
High Elo	0.083 (<0.001)	0.0702 (0.003)		
	(0.019)	(0.0205)		
H Adviser×Known Adviser	-0.168 (<0.001)	-0.1685 (<0.001)	-1.7e-01 (<0.001)	-1.7e-01 (<0.001)
	(0.038)	(0.0380)	(3.8e-02)	(0.03812)
Elo			1.7e-04 (0.010)	2.2e-04 (0.002)
			(6.1e-05)	(0.00006)
Elo×rated			4.7e-06 (0.881)	-1.2e-05 (0.685)
			(3.1e-05)	(0.00003)
Control Index		0.0331 (0.090)		3.4e-02 (0.087)
		(0.0185)		(0.01901)
Male		0.0562 (0.092)		4.3e-02 (0.163)
		(0.0316)		(0.02957)
Age		0.0012 (0.267)		9.2e-04 (0.396)
		(0.0011)		(0.00107)
Std.Errors	by: position	by: position	by: position	by: position
Num.Obs.	1999	1749	1999	1749
R2	0.064	0.069	0.073	0.080
R2 Adj.	0.052	0.055	0.061	0.065

Notes: Robust standard errors clustered at the position level. In parenthesis on the same line are the p-value, below the standard error.

(3) and (4) add demographic controls but restrict the sample.

[1] Absolute distance from the correct answer in pawn advantage.

B.2 Correct Answers

In Table 9, we see that high ability subjects have higher rate of correct answers than low ability ones. The only exception is the Elo heuristic, because in the Unknown Adviser treatment, h type subjects are penalized by never following the H advice at all. The probability heuristic tells us that they should sometimes follow the mixed advice. The regression in Table 8 shows that higher Elo subjects tend to keep their answer more.. Importantly, to be followed by subjects, the H advice has to be revealed as one, as shown by the interaction term beteen H adviser and Known Adviser, compared to the

insignificant H adviser term alone.

Table 9: Difference (in percentage points) between the average share of correct answer of h and l type subjects, before advice, post advice, and following our two heuristics.

Treatment	Before	After	Probabilistic	Elo
All	10.3	9.4	6.5	2.2
Unknown Adviser	10.7	7.6	7.0	-3.5
Known Adviser	10.0	10.9	5.8	6.6

The regression in Table 10 adds to the regression in Table in 5 demographic controls. The results are consistent between the two and the additional control variables are not significant.

Table 10: Regression for the share of correct answers, with fixed effects at the position level and controlling for individual characteristics.

	Grouped by Elo		Continous Elo	
	Known	Unknown	Known	Unknown
H Adviser	0.1533 (0.020)	0.0701 (0.225)	0.31017 (0.050)	0.34626 (0.048)
	(0.0606)	(0.0559)	(1.5e-01)	(0.16387)
Low Elo	-0.1656 (<0.001)	-0.1366 (0.051)		
	(0.0346)	(0.0657)		
Elo			0.00030 (<0.001)	0.00026 (0.029)
			(5.9e-05)	(0.00011)
H Adviser×Low Elo	0.0460 (0.368)	0.0924 (0.146)		
	(0.0500)	(0.0610)		
H Adviser×Elo			-0.00011 (0.311)	-0.00018 (0.167)
			(1.0e-04)	(0.00013)
Male	-0.0473 (0.186)	-0.0869 (0.107)	-0.07296 (0.047)	-0.09827 (0.068)
	(0.0345)	(0.0514)	(3.4e-02)	(0.05081)
Age	-0.0036 (0.053)	-0.0035 (0.072)	-0.00259 (0.191)	-0.00360 (0.077)
	(0.0018)	(0.0019)	(1.9e-03)	(0.00193)
Control Index	0.0169 (0.601)	0.0576 (0.165)	0.00247 (0.939)	0.05682 (0.167)
	(0.0317)	(0.0399)	(3.2e-02)	(0.03953)
Std.Errors	by: position	by: position	by: position	by: position
Num.Obs.	960	800	960	800
R2	0.168	0.086	0.168	0.091
R2 Adj.	0.145	0.056	0.146	0.061

Robust standard errors clustered at the position level.

In parenthesis on the same line are the p-value, below the standard error.

C Experimental Instructions (Online Appendix)

FOR ONLINE PUBLICATION ONLY: EXPERIMENTAL INSTRUCTIONS IN THE KNOWN ADVISOR TREATMENT

Response Sheet - ورقة الاجابة

Participant number – (ة)رقم المشترك:A1_____

ELO - التصنيف:

Your predictions - التوقعات

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								

Round 2 – الجولة الثانية:

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								

(Please complete both sides of the sheet)

الرجاء تعبئة جانبي الورقة

Additional Info – معلومات إضافية

Age - العمر:

Gender - الجنس:

How much do you agree with the following statements - ما مدى موافقتك على العبارات التالية ؟

	Strongly disagree أعارض بشدة	Disagree أعارض	Neither agree nor disagree لا أوافق ولا أعارض	Agree أوافق	Strongly agree أوافق بشدة
I try to avoid situations where someone else tells me what to do. أحاول تجنب المواقف التي يقول لي فيها شخص آخر بما يجب القيام به.					
I prefer to be a leader rather than a follower. أفضل أن أكون قائدًا وليس تابعًا.					
I enjoy making my own decisions. أنا أستمتع باتخاذ قراراتي بنفسي.					
I would rather someone else took over the leadership role when I'm involved in a group project. أفضل أن يتولى شخص آخر الدور القيادي عندما أشارك في مشروع جماعي.					
There are many situations in which I would prefer only one choice rather than having to make a decision. هناك العديد من المواقف التي أفضل فيها خيارًا واحدًا فقط بدلاً من الاضطرار إلى اتخاذ قرار.					

A1i

Here are ten positions that occurred in real chess games which have been chosen from a dataset of previous games from the Mega Database 2023.

We will ask you to evaluate 20 games over two rounds: 1 and 2. We will pick one of your evaluations at random and you will receive a voucher of $10 if your answer was correct.

Please complete **Round 1, Part 1** of the Response Sheet by indicating for each game your best estimate of the pawn advantage, which can be +0.7, -0.7, +2.4, or -2.4. Please check the box corresponding to your choice (only one possible answer). Note that the positions have a pawn advantage of ±0.7 and one of ±2.4 with equal probability.

Once you have completed **Round 1, Part 1**, please wait for the experimenter to give you the next set of instructions.

You have a total of 8 minutes to complete this part.

فيما يلي عشرة أوضاع حصلت في جولات شطرنج حقيقية وقد تم اختيارها من مجموعة بيانات للألعاب السابقة من Mega Database 2023.

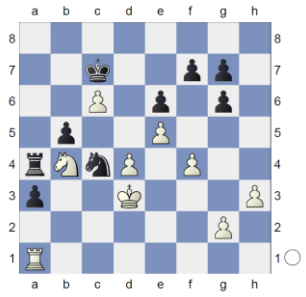
سنطلب منك تقييم 20 وضع على مرحلتين: الجولة الأولى والجولة الثانية. سوف نختار أحد تقييماتك بشكل عشوائي وستتلقى قسيمة بقيمة 10 دولارات إذا كانت إجابتك صحيحة.

يرجى تعبئة **الجولة 1، الجزء 1** من ورقة الإجابة بالإشارة إلى أفضل تقدير لديك لكل وضع حسب أفضلية ال pawn advantage والتي يمكن أن تكون 0.7+ أو 0.7- أو 2.4+ أو 2.4-. يرجى تحديد المربع المقابل لاختيارك (إجابة واحدة فقط ممكنة). ملاحظة: من المحتمل أن يكون الوضع مع أفضلية 0.7±، **أو** 2.4±، مع احتمالية متساوية.

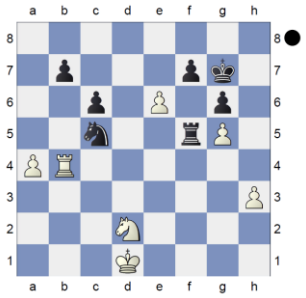
بمجرد الانتهاء من الجولة 1، الجزء 1، من فضلك انتظر أن يعطيك المشرف المجموعة التالية من التعليمات.

لديك 8 دقائق لإكمال هذا الجزء.

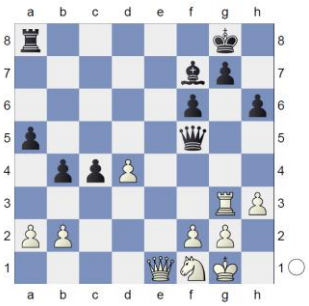
1



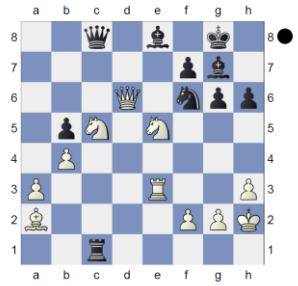
2



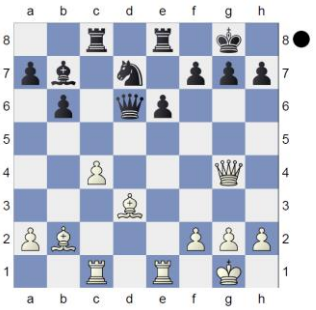
3



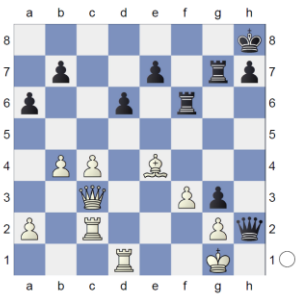
4



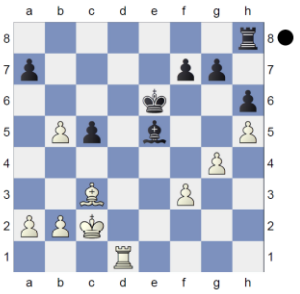
5



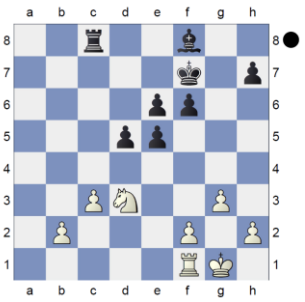
6



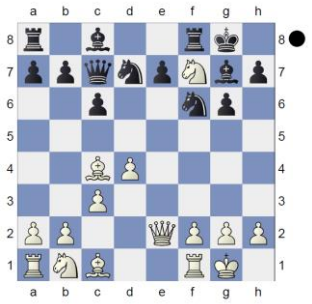
7



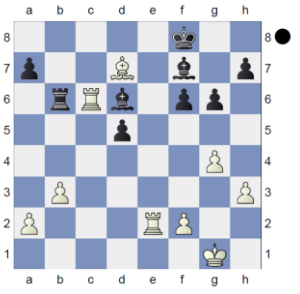
8



9



10



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.

سنزودك الآن ببعض المعلومات الإضافية حول الاوضاع العشر.

لقد طلبنا من **لاعب(ة) تصنيفه 2335** أن يقيم الاوضاع العشر في نفس ظروفك. يمكنك الاطلاع على توقعاتهم في الجدول أدناه.

بالنظر إلى تقديرك في **الجولة 1، الجزء 1** في ورقة الإجابة، يرجى إكمال **الجولة 1، الجزء 2**. لك مطلق الحرية في تغيير توقعاتك السابقة أو الاحتفاظ بها بناءً على المعلومات الواردة في هذه الورقة.

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.

A1ii

Now, we will repeat the previous exercise with a new set of ten positions.

Please complete **Round 2, Part 1** of the Response Sheet. This is the same procedure as for **Round 1.**

Once you have completed **Round 2, Part 1**, please wait for the experimenter to give you the next set of instructions.

الآن، سنكرر التمرين السابق بمجموعة جديدة من عشر أوضاع.

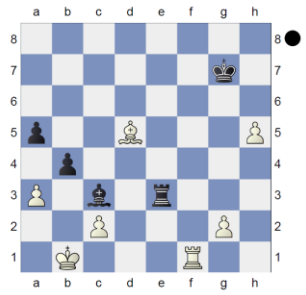
يرجى تعبئة **الجولة الثانية، الجزء 1** من ورقة الإجابة. هذا هو نفس الإجراء المتبع في الجولة الأولى.

بمجرد الانتهاء من الجولة 2، الجزء 1، من فضلك انتظر أن يعطيك المشرف المجموعة التالية من التعليمات.

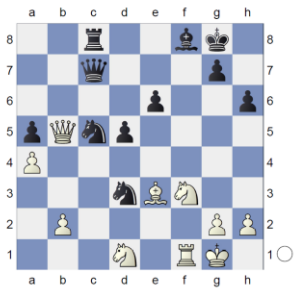
You have a total of 8 minutes to complete this part.

لديك 8 دقائق لإكمال هذا الجزء.

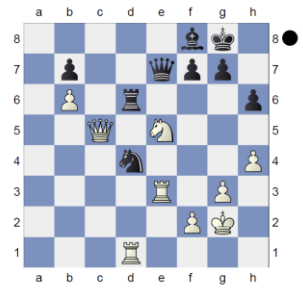
11



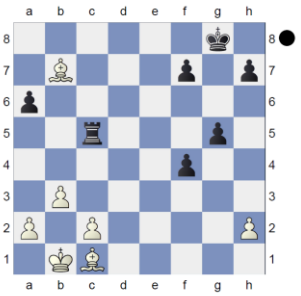
16



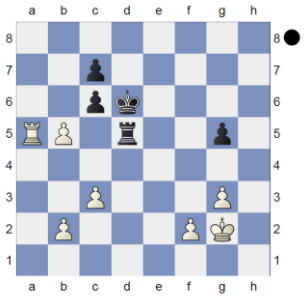
12



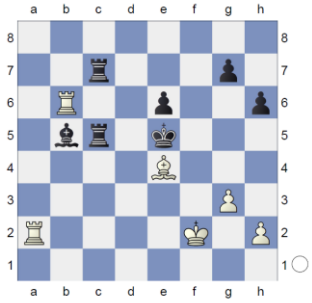
17



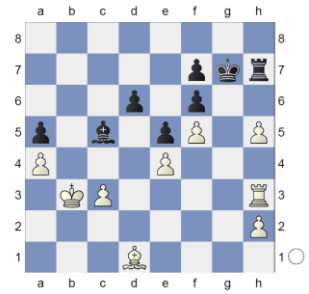
13



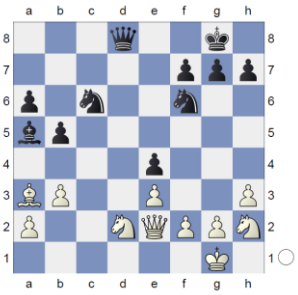
18



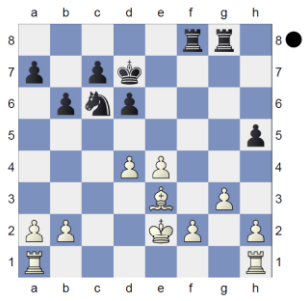
14



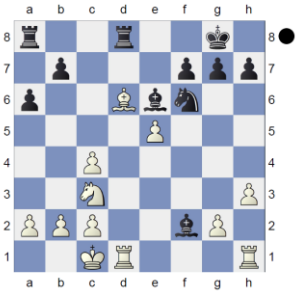
19



15



20



A1ii

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

We have asked **an unrated player**, who plays regularly for fun, to evaluate the ten positions in the same conditions as you. You can find their predictions in the table below.

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

سنزودك الآن ببعض المعلومات الإضافية حول الاوضاع العشر.

لقد طلبنا من **لاعب(ة)غير مصنف، يلعب بانتظام من أجل التسلية،** أن يقيم الاوضاع العشر في نفس ظروفك. يمكنك الاطلاع على توقعاتهم في الجدول أدناه.

بالنظر إلى تقديرك في **الجولة 2، الجزء 1** في ورقة الإجابة، يرجى إكمال **الجولة 2، الجزء 2**. لك مطلق الحرية في تغيير توقعاتك السابقة أو الاحتفاظ بها بناءً على المعلومات الواردة في هذه الورقة.

رقم الوضع Position Number	التفوق Pawn advantage
11	-2.4
12	+0.7
13	+2.4
14	-0.7
15	+0.7
16	-2.4
17	-0.7
18	-2.4
19	-0.7
20	-0.7

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

لديك 4 دقائق لإكمال هذا الجزء.

When this is over, please complete the personal information questions at the back of the response sheet.

عندما تنتهي من هذا الجزء، يرجى إكمال أسئلة المعلومات الشخصية في الجزء الخلفي من ورقة الإجابة.