Supporting Information for "Personal Bests as Reference Points"

Ashton Anderson and Etan A. Green

Opponent selection

Players choose opponents, and those choices entail risk. Do reference-dependent risk preferences predict opponent selection?

Theoretically, a player's choice of opponent—and hence, her choice of her reward for winning and corresponding penalty for losing—follows from her risk preferences. To see this, first consider a risk-averse player, who only tolerates an uncertain outcome if the expected gain is sufficiently positive. A key feature of the rating system is that in any game, a player's expected rating change is zero—regardless of her opponent's rating. Players win less frequently against higher-rated opponents, but the reward for winning and the penalty for losing are calibrated to be perfectly compensating. (The rating system is further described in the Materials and Methods.) Since the expected rating change of any game is always zero, risk-averse players choose extreme victory rewards of either 0 points, effectively guaranteeing a win against an overmatched opponent, or 16 points, effectively guaranteeing a loss against a far stronger opponent; in both cases, ratings remain the same. As a result, risk-seeking preferences are most relevant for our context, since they rationalize a willingness to play opponents against whom the outcome is uncertain.

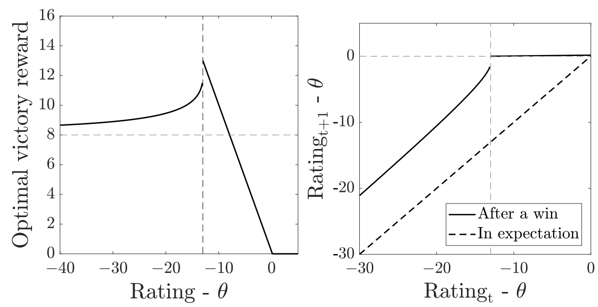
We begin by estimating theoretically optimal victory rewards given preferences that define risk tolerance in relation to a reference point θ (1):

$$v_\theta(x) = \begin{cases} x^\alpha & x > \theta \\ -\lambda(-x)^\alpha & x \leq \theta \end{cases}$$

where $\lambda > 1$ creates a kink at the reference point, and $0 < \alpha < 1$ produces convex utility in losses and concave utility in gains. Below the reference point, players are risk seeking and willing to gamble with their ratings; above the reference point, players are risk averse and unwilling to gamble.

We first consider the optimal victory reward for a reference-dependent player who evaluates her rating after the next game. This choice is akin to that of a gambler who places a bet with the goal of getting back to even, or to the chess player who chooses a victory reward with the goal of eclipsing her personal-best rating. We then consider the optimal victory reward for a reference-dependent player who evaluates her rating both after the next game and after a subsequent game. Finally, we consider the optimal victory reward under an infinite horizon.

Figure S1a shows the optimal victory reward under a one-game horizon as a function of the player's rating distance from the reference rating. When a player's rating exceeds the reference rating, she chooses a victory reward of 0 (or equivalently, 16), which fixes her rating in place. Short of the reference point, players choose victory rewards that target the reference point. In Figure S1a, the linear relationship for players who are within about 12 points of the reference rating has a slope of -1, implying that the optimal victory reward equals the distance between her rating and the reference point. We show this directly in Figure S1b, which compares a player's current rating to her rating after choosing the optimal victory reward, both in expectation and conditional on winning. All gambles preserve a player's current rating in expectation, but different gambles have different upsides. Short of the reference point, the reference-dependent player selects the opponent against whom a win would vault her rating to the goal—and no further.

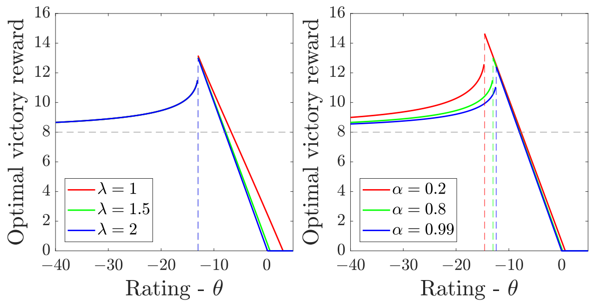


(a) Optimal victory reward **(b)** Next rating under optimal reward

Fig. S1. One-game horizon ($\lambda = 2, \alpha = 0.8$).

At some distance this strategy becomes untenable, given that a player who chooses a victory reward of 16 will never win. Hence, there exists a lesser distance beyond which the player stops trying to achieve her goal in a single bound. For the parameter values underpinning Figure S1, this point occurs when a player is about 12 points from the reference rating. Beyond that distance, eclipsing the reference point in a single shot is sufficiently unlikely, and the optimal victory reward drops discontinuously and then asymptotes to 8, implying an equally rated opponent. In effect, the player 12 or more points from her best-ever rating gives up on eclipsing that reference point in her next game.

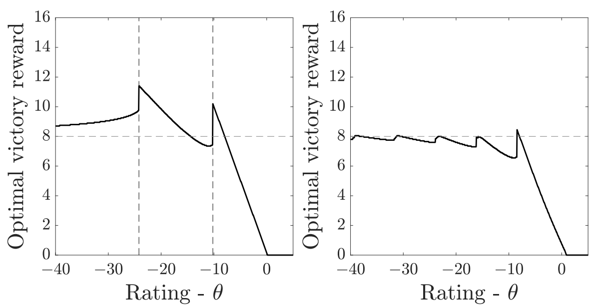
This pattern holds for a range of parameter values. Figure S2 shows the optimal victory reward function for different values of λ (S2a) and α (S2b). Increasing λ



(a) $\alpha = 0.8$ and $\lambda \in \{1, 1.5, 2\}$ **(b)** $\alpha \in \{0.2, 0.8, 0.99\}$ and $\lambda = 2$

Fig. S2. Optimal victory reward in one-game horizon, varying λ and α.

increases loss aversion, which brings the target rating closer to θ. Whereas loss-neutral (i.e., $\lambda = 1$) actors target ratings that eclipse the reference point, even mildly loss-averse actors target the reference point almost exactly. Loss aversion penalizes outcomes that fall short of the reference point. Hence, loss-averse actors are unwilling to gamble on the upside risk of eclipsing the reference point because doing so entails more downside risk of falling short. Separately, decreasing α makes the player more risk seeking in losses, thereby increasing the peak optimal victory reward by increasing the player's tolerance for matches that she will almost surely lose.



(a) Two-game horizon **(b)** Infinite horizon

Fig. S3. Optimal victory reward ($\lambda = 2, \alpha = 0.8, \delta = 0.99$).

So far, we have considered only a one-game horizon. In chess, however, players typically play a series of games—choosing an opponent, experiencing the outcome, and then choosing another opponent. To capture this dynamic, we consider the optimal victory reward for players who evaluate their ratings not only after the upcoming game (as shown in Figure S1a) but after subsequent games as well, as shown in Figure S3. In particular, we evaluate the optimal victory reward function using dynamic programming under a two-game horizon (S3a) and under an infinite horizon (S3b). Here, the outcome of the later games is subject to a discount factor $0 < \delta < 1$, which implies that outcomes farther in the future have less bearing on valuations in the moment.

Under a two-game horizon, as under a one-game horizon, a player sufficiently close to the reference point chooses the opponent against whom a victory would vault her rating to the goal. Likewise, the player gives up on her goal when she is sufficiently far away. But now, an in-between region emerges in which the player can reasonably achieve her goal in two steps. In this region, the player does not target the reference point in her first game. Rather, she positions herself to reach the reference point after a second game.

As the horizon approaches infinity (S3b), the non-monotonic pattern from the two-game horizon repeats and asymptotes to a victory reward of 8 with distance from the reference point. With many opportunities to reach the reference point, players play weaker opponents. Under a one-game horizon, the optimal victory reward peaks 13.0 points short of the reference point for $\alpha = 0.8$ and $\lambda = 2$. For these same parameter values, and a discount factor of $\delta = 0.99$, the optimal victory reward peaks 11.4 points short of the reference point under a two-game horizon, and it peaks just 8.5 points short of the reference point under an infinite horizon.

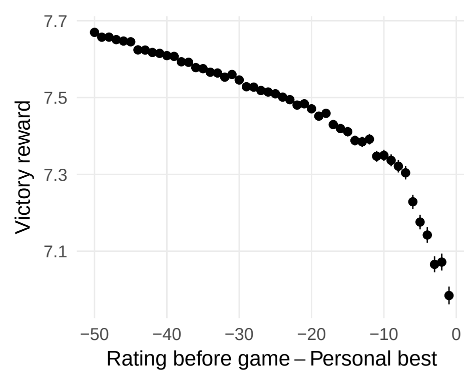


Fig. S4. Average observed victory reward, with 95% confidence intervals.

Does the optimal reward function under a one-game, two-game, or infinite horizon predict observed victory rewards in the data? Figure S4 shows the average observed victory reward as a function of the player's rating distance from her personal best. The pattern is monotonically decreasing in a manner that is inconsistent with reference-dependent opponent selection under any time horizon.

The downward trend is an artifact produced by Simpson's paradox. When restricted to player-games following a win or a loss, the trend is relatively flat following a win and increasing following a loss. However, the levels differ: victory rewards are higher following a loss than a win—for example, players who lose are more likely to have played higher-rated opponents and, because rematches are prevalent, to subsequently play them again. Further, the relative weighting between these two groups changes with proximity to the reference point. Far from one's personal best, players are equally likely to have arrived by way of a win or a loss. But close to the reference point, players are more likely to have arrived by way of a win. For example, the only way to be 1 rating point short of the reference point after a loss is to have previously been exactly at the reference point and lose against a vastly higher-rated opponent. This differential weighting produces the monotonically decreasing pattern when games following a win are pooled with games following a loss.

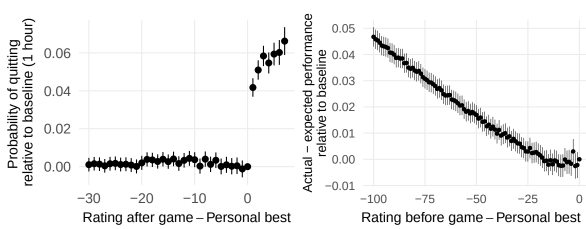
Robustness Checks

A. Within-player estimates. Player heterogeneity potentially confounds our estimates. If players who are close to their personal-best ratings have different baseline outcomes than those farther from their personal-best ratings, then average outcomes will differ with distance to the reference point because the players differ, not necessarily because behavior changes as players approach their personal-best ratings. Note that this is not a concern when measuring discontinuous changes at a boundary, under the assumption that the player population changes smoothly across the boundary.

To address this potential confound, we estimate a regression with player fixed effects:

$$y_{p,g} = \alpha_p + \sum_i \mathbb{1}\{\text{Rating} - \text{personal best} = i\} \cdot \beta_i + \epsilon_{p,g}, \quad [1]$$

where $y_{p,g}$ is an outcome (i.e., quitting or performance) for player p and player-game g, α_p are player fixed effects, β_i are the parameters of interest, and $\epsilon_{p,g}$ is an error term. The player fixed effects α_p allow for player-specific intercepts, implying that the coefficients β_i measure the within-player change in y when the difference between a player's rating and her personal best equals i.



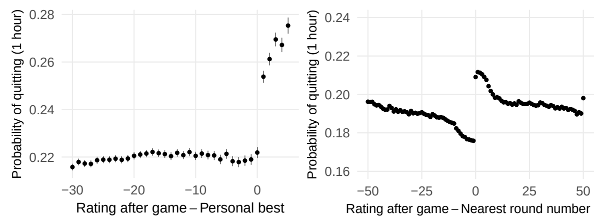
(a) Quitting. **(b)** Performance.

Fig. S5. Estimates with player fixed effects ($\hat{\beta}_i$ from Equation 1).

Figure S5 shows within-player changes in the 1-hour quitting rate (S5a) and in performance minus expectations (S5b). For both figures, the outcome when a player's rating equals her personal best is normalized to 0, and the 95% confidence intervals are calculated from standard errors clustered by player.

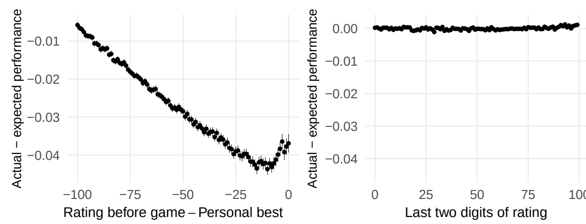
The figures are similar to their counterparts in the main text. With player fixed effects, the probability of quitting jumps 4.2 percentage points after a player sets a personal best, compared to 4.5 percentage points without fixed effects. With and without player fixed effects, the prevailing regression-to-the-mean trend abates when the player's rating is near to her personal best.

B. Personal best recency restriction. For the main analyses, we exclude player-games in which the player set a personal best in her last 20 games. Here, we reproduce our two main empirical figures without this restriction, expanding the sample to 220M player-games across 142M unique games. As can be seen in Figures S6 and S7, the results are qualitatively identical to those presented in the main text.



(a) Personal bests. **(b)** Round numbers.

Fig. S6. Probability of quitting for at least 1 hour absent the recency restriction, with 95% confidence intervals.

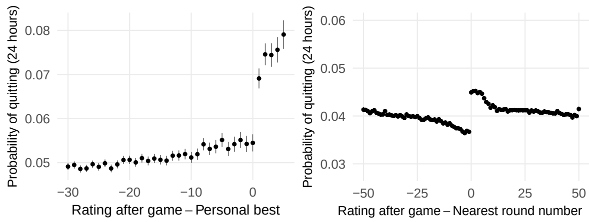


(a) Personal bests. **(b)** Round numbers.

Fig. S7. Performance short of personal bests and round numbers absent the recency restriction, with 95% confidence intervals.

C. Definition of quitting. We define quitting as not playing a blitz game for at least one hour. However, the results are qualitatively identical for a longer threshold. Figure S8 shows the probability of quitting when quitting is defined as not playing a blitz game for 24 hours. As ratings eclipse personal bests (S8a), the probability of quitting jumps 1.5 percentage point, or 27.7%. As ratings eclipse multiples of 100 (S8b), the probability of quitting jumps 0.6 percentage points, or 15.5%.

D. Effort after setting a new personal best. Figure S9 measures performance on the subset of player-games that follow a win, which allows us to compare players who are approaching their personal bests to those who just set a new personal best by winning the previous game. As ratings approach the reference point, the prevailing regression-to-the-mean trend flattens, consistent with prior results. Moreover, performance does not appear



(a) Personal bests.

(b) Round numbers.

Fig. S8. Probability of quitting for at least 24 hours, with 95% confidence intervals.

to regress after setting a new personal best, suggesting that players continue exerting effort to set another personal best.

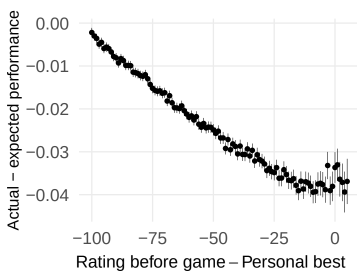


Fig. S9. Performance following a win.

E. Calibration of rating system. Figure S10 shows a calibration plot, comparing predicted and actual win rates at each victory reward when $k = 16$. Predicted and actual win rates match closely at every victory reward. The actual win rate is slightly higher than the predicted win rate when the victory reward is 15, but removing player-games where the victory reward is 15 (or all player-games where the victory reward exceeds 12) does not substantively alter the results.

Fig. S10. Predicted (line) and actual (dots) win rates at each victory reward.

1. D Kahneman and A Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.