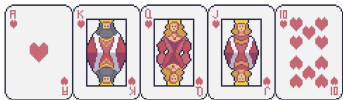


Table S1. The ranking of Poker hands in common variants like Texas Hold 'Em, with rank 1 (royal flush) representing the best possible hand.

Rank	Type	Description	Example
1	Royal flush	A, K, Q, J, 10, all with the same suit	
2	Straight flush	Any straight with the same suit	
3	Four of a kind	Four cards of equal value with any other card	
4	Full house	Three cards of equal value with any pair	
5	Flush	Any five cards with the same suit	
6	Straight	Five cards with consecutive values but not all of the same suit	
7	Three of a kind	Three cards with the same value and two other cards with different values	
8	Two pairs	Two pairs consist of two sets of two cards with the same value, along with one additional card	
9	Pair	One pair consist of two cards with the same value, along with three additional cards	
10	High card	Five cards that do not form any of the above-mentioned hands	

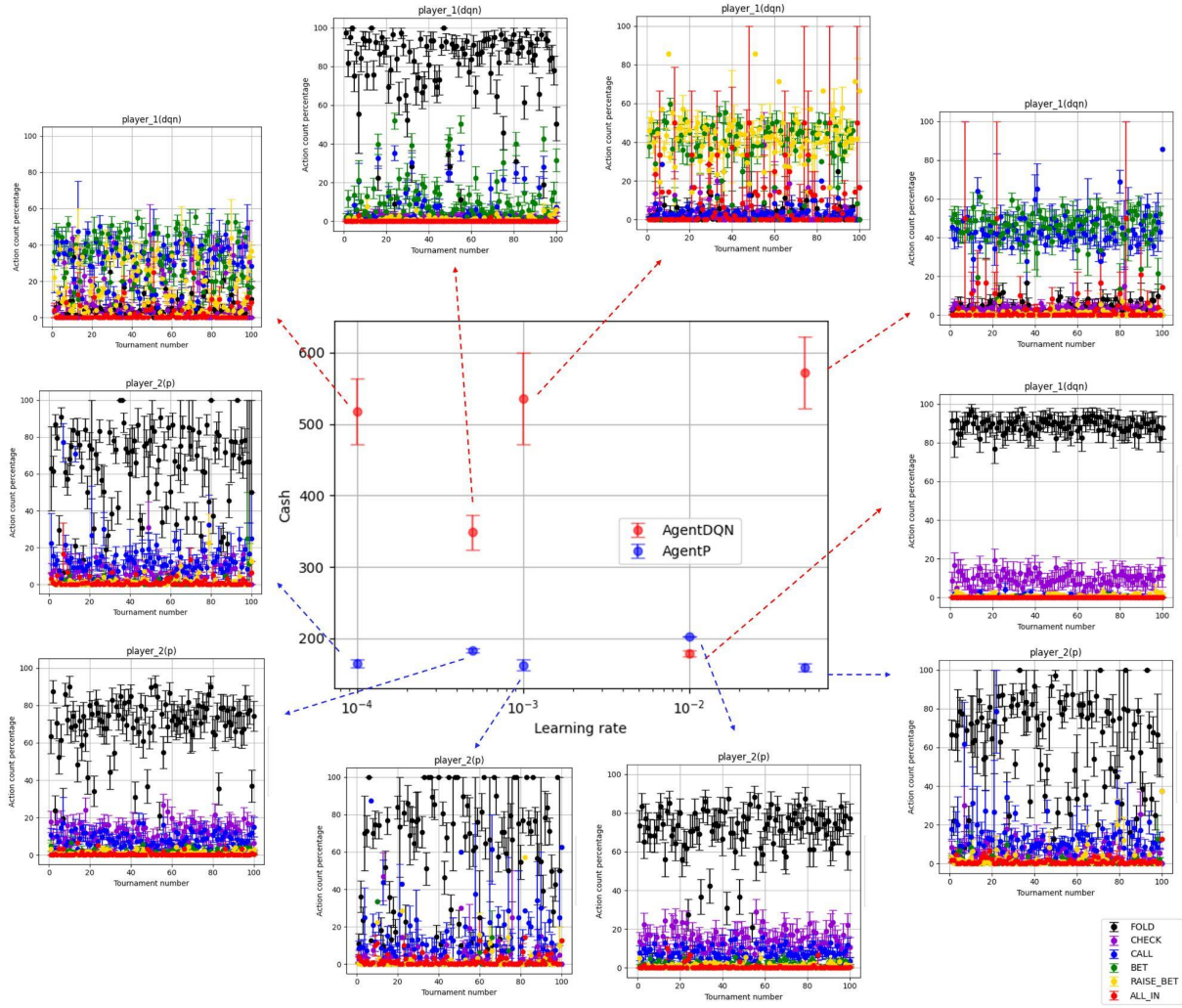


Figure S1. The performance and action profile of AgentDQN at different learning rates. Each blue point represents AgentP's performance when competing against AgentDQN at a specific learning rate. They are plotted on the same x-axis for better comparison. For each data point, the corresponding action profile is also plotted. AgentP performs consistent actions, most likely folding. AgentDQN exhibits more complex actions at a learning rate of 10^{-4} , in contrast to other learning rates where it performs more monotonous actions.

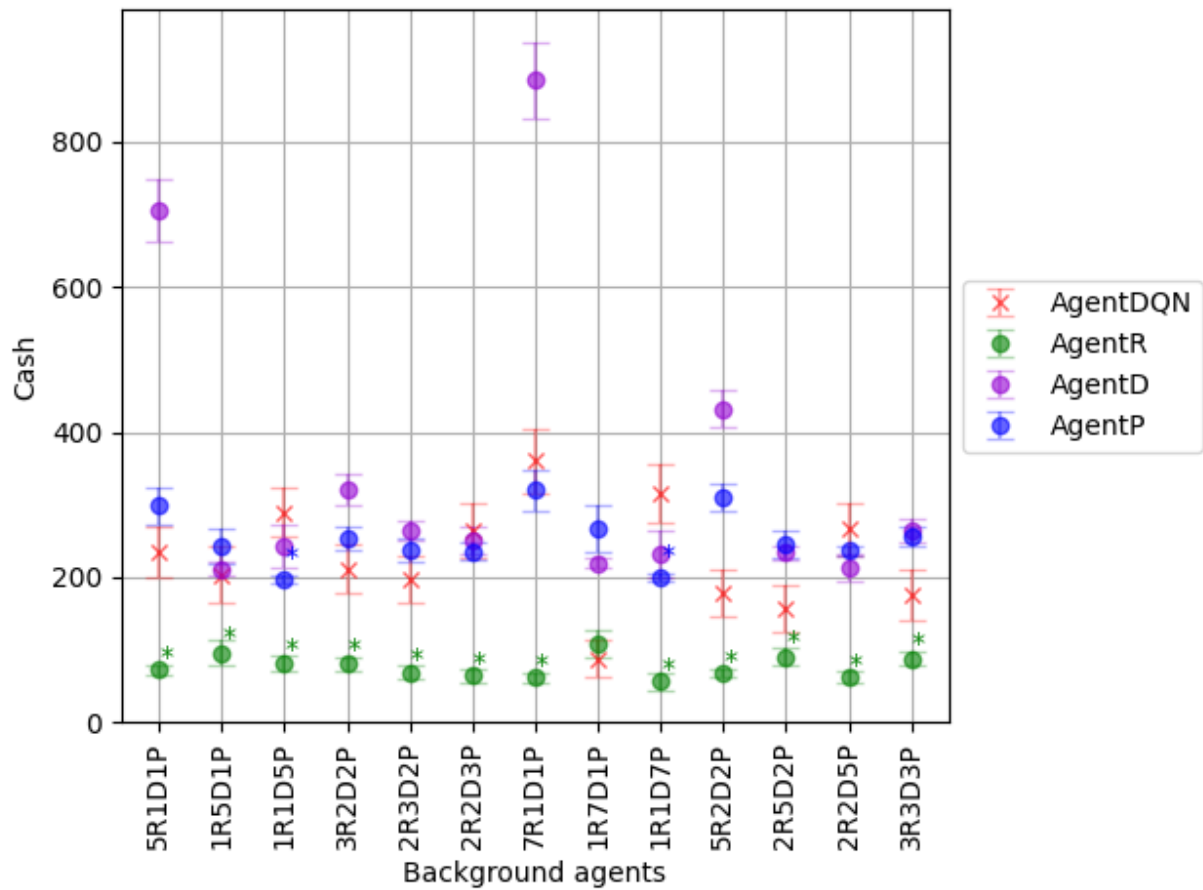


Figure S2. Performance of four agents (measured by the cash balance metric) when AgentDQN competes against different combinations of background agents. AgentDQN was trained with nine AgentP, leading to slightly better performance than other agents in the combinations dominated by AgentP, such as 1R1D5P, 2R2D3P 1R1D7P, and 2R2D5P. Similar trends are observed when competing against single-type agents: AgentDQN significantly outperforms AgentR but performs similarly to AgentD.

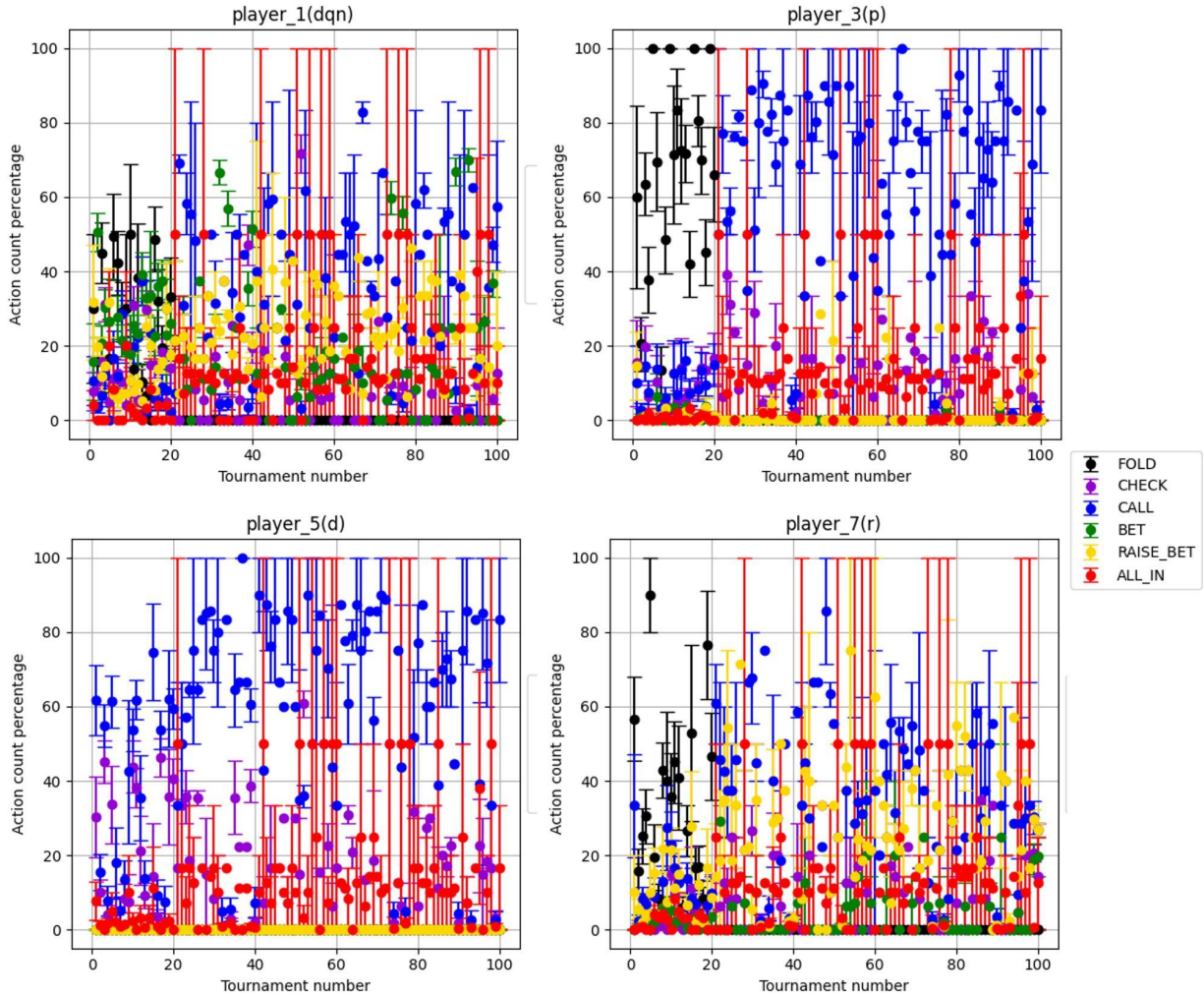


Figure S3. The action profile of selected four players in an environment with the novelty GameFoldRestrict, where AgentDQN competes against three AgentP, two AgentD, and two AgentR. The average cash balance of AgentDQN dropped from 235.12 ± 69.58 to 214.0 ± 54.44 from pre-novelty to post-novelty, while AgentP's performance changed from 242.63 ± 55.03 to 197.20 ± 51.92 , AgentD's from 278.39 ± 94.97 to 197.19 ± 54.54 , and AgentR's performance changed from 39.62 ± 9.72 to 200 ± 53.5 . As described in the main text, the fold is prohibited in the post-novelty tournaments. Interestingly, AgentR, which performs random actions, outperforms the other three types of agents in the post-novelty tournaments, indicating that the original "knowledge or strategies" are not applicable to this novel situation.

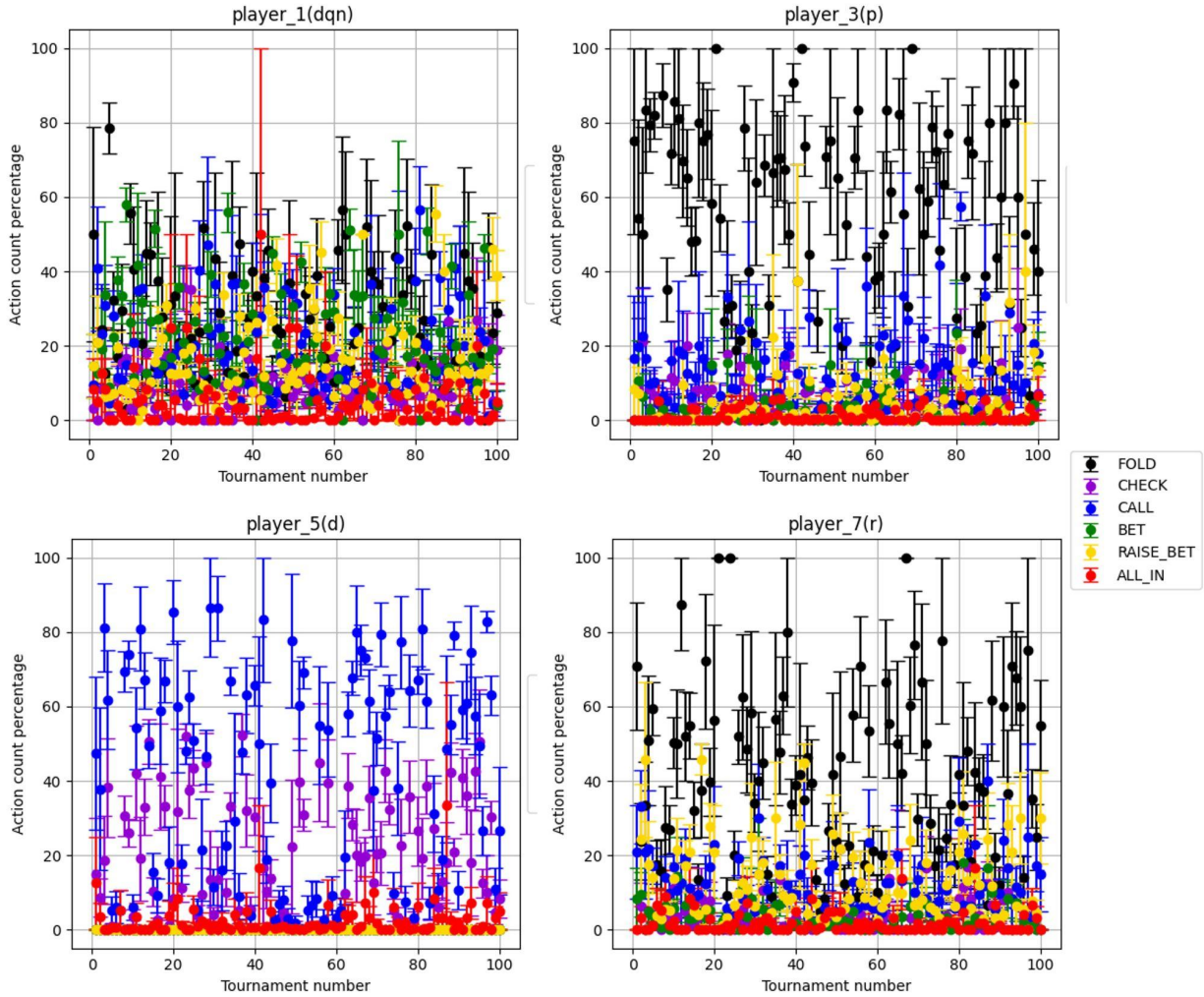


Figure S4. The action profile of selected four players in an environment with the novelty CardDistHigh, where AgentDQN competes against three AgentP, two AgentD, and two AgentR. The average cash balance of AgentDQN dropped from 287.5 ± 86.01 to 134.65 ± 28.29 from pre-novelty to post-novelty, while AgentP's performance changed from 203.92 ± 37.45 to 295.29 ± 41.34 , AgentD's from 259.50 ± 83.95 to 226.78 ± 43.42 , and AgentR's performance changed from 89.5 ± 30.63 to 62.48 ± 10.74 . This novelty alter card distribution to the high-value cards, which benefits AgentP as it strategizes to bet on good hands. The corresponding actions are also shown in the top-right sub-figure, where we observe an increase in raise frequency in the post-novelty tournaments.

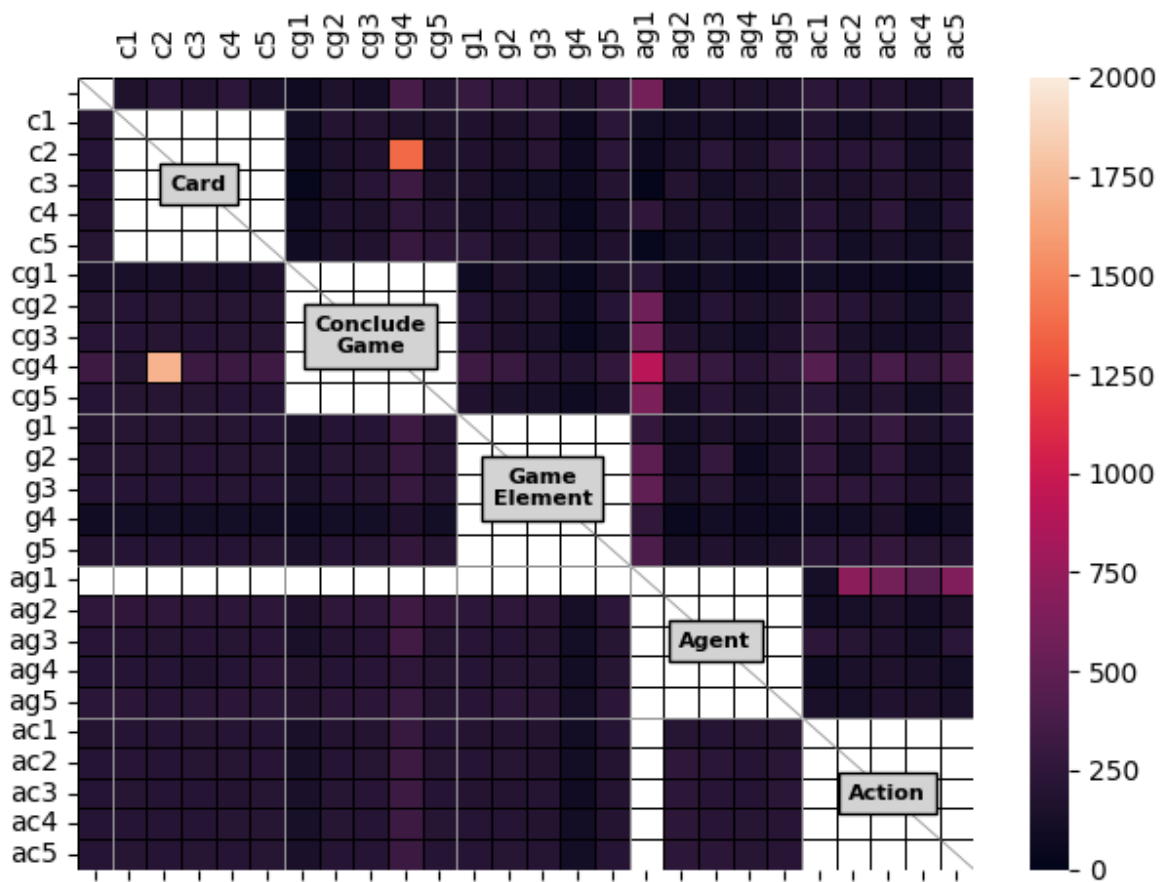


Figure S5. Performance of AgentDQN against AgentD when encountering depth-two novelties.

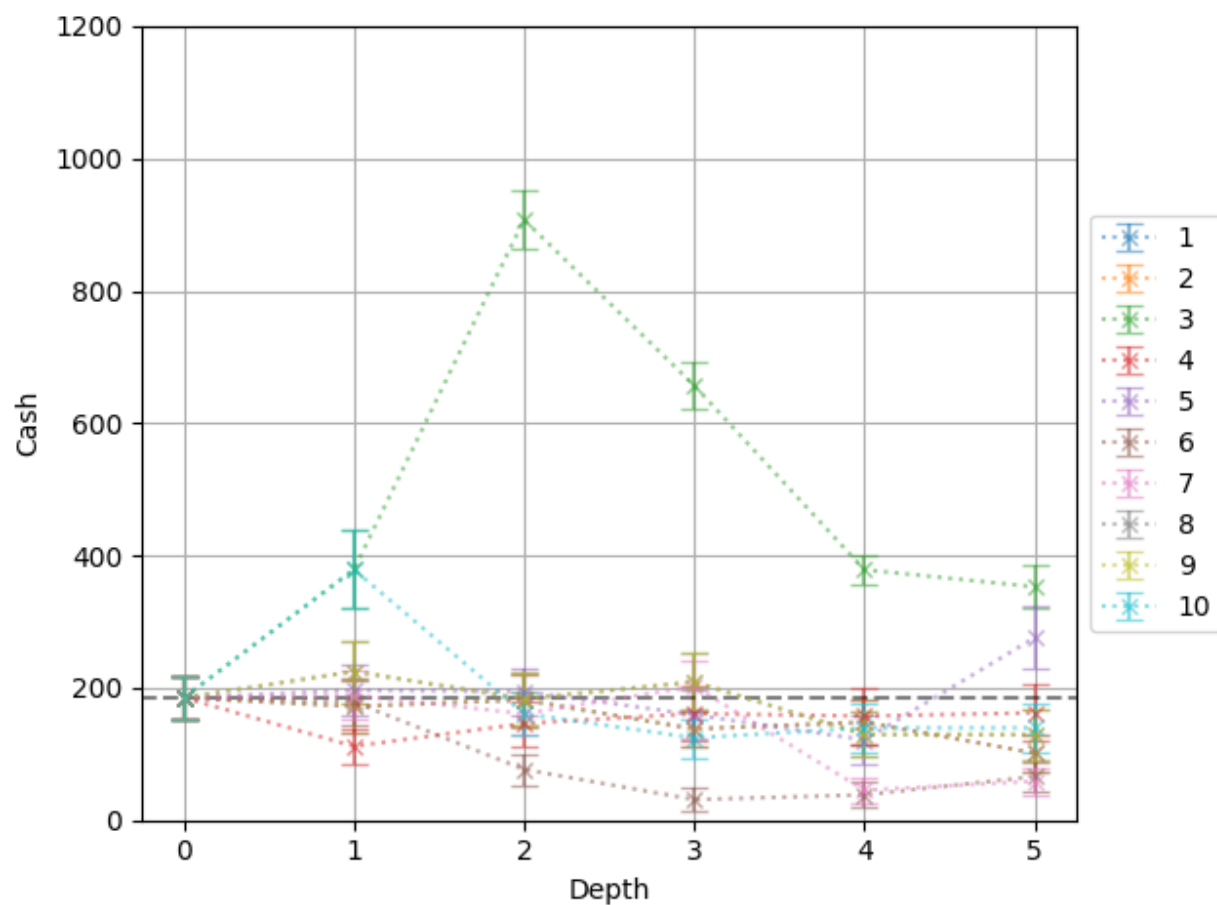


Figure S6. Performance of AgentDQN against AgentD in randomly selected composite novelties with depths ranging from 0 to 5, where depth 0 represents the original setting.

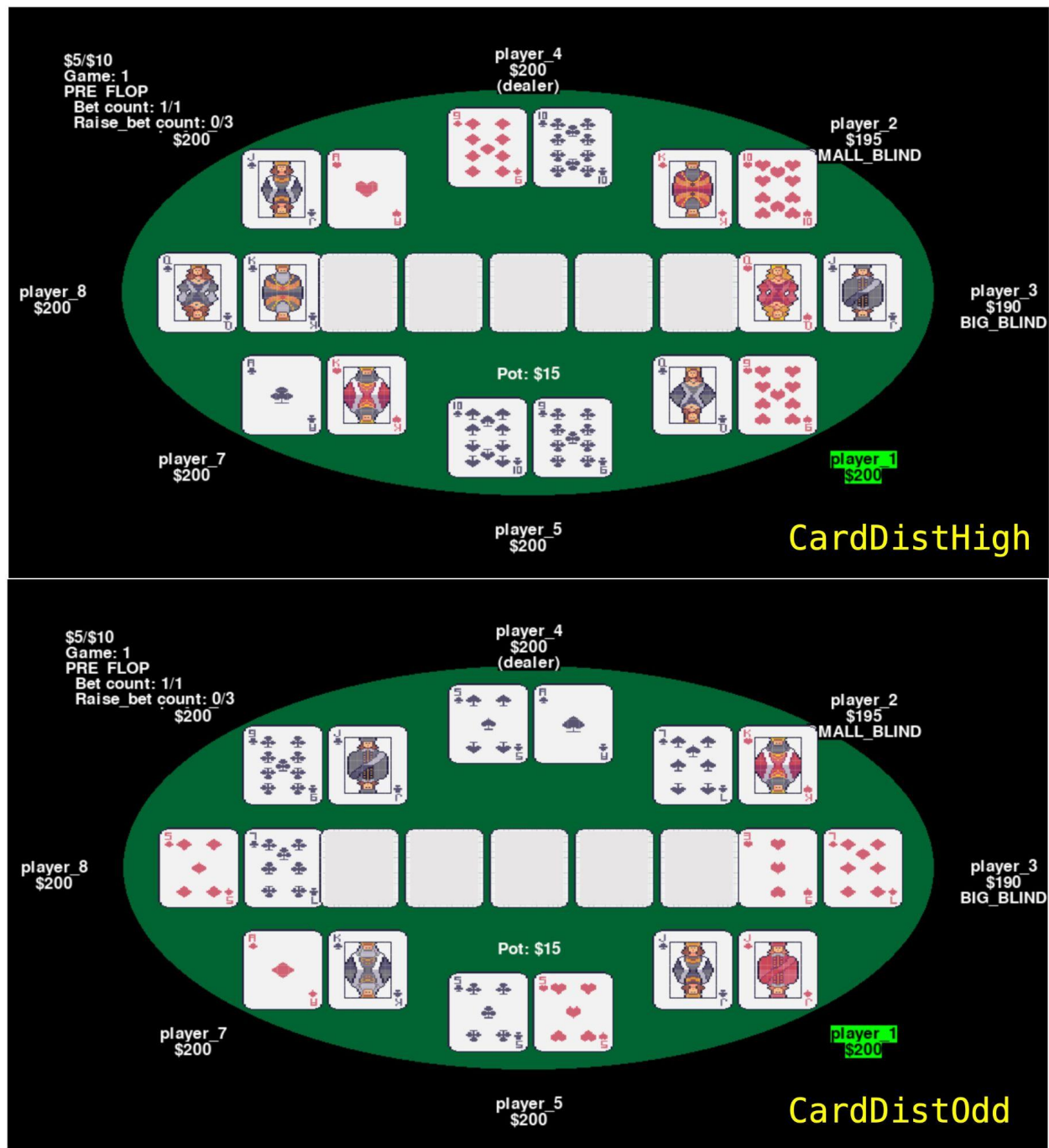


Figure S7. Two selected novelties in debug visualization mode. Top: CardDistHigh, altering the card distribution by placing the high-number cards last in the deck. This novelty affects players' hole card strength. For example, a pair of Queens is no longer relatively strong. Bottom: CardDistOdd, altering the card distribution by positioning a particular parity. In this novelty, straights will no longer be possible, thus significantly changing the strategy players must use.