¹ Transfer of Learned Opponent Models in Zero Sum Games

² Ismail Guennouni[1] & Maarten Speekenbrink[1]

³ [1] Department of Experimental Psychology, University College London

⁴ Author Note

⁵ Add complete departmental affiliations for each author here. Each new line herein

⁶ must be indented, like this line.

⁷ Enter author note here.

⁸ Correspondence concerning this article should be addressed to Ismail Guennouni,

⁹ Department of Experimental Psychology, University College London, 26 Bedford Way,

¹⁰ London WC1H 0AP, United Kingdom. E-mail: i.guennouni.17@ucl.ac.uk

11   Abstract

12   Enter abstract here. Each new line herein must be indented, like this line.

13   *Keywords:* keywords

14   Word count: X

Transfer of Learned Opponent Models in Zero Sum Games

Introduction

Being able to transfer previously acquired knowledge to a new domain is one of the hallmarks of human intelligence (Taylor et al., 2008). Humans are naturally endowed with the ability to extract relevant features from a situation, identify the presence of these features in a novel setting and use previously acquired knowledge to adapt to previously unseen challenges using acquired knowledge. More formally, Perkins (1992) defines transfer of learning as the application of skills, knowledge, and/or attitudes that were learned in one situation to another learning situation. This typically human skill has so far eluded modern ai agents. Deep neural networks for instance can do very well on image recognition tasks and can even reach super-human performance levels on video and strategic board games. Yet they struggle to learn as fast or as efficiently as humans do, and more importantly they have a very limited ability to generalize and transfer knowledge to new domains. Lake et al.(2017) argue that human learning transfer abilities take advantage of important cognitive building blocks such as a deeper representation of concepts underlying tasks and compositionally structured causal models of the environment.

In this paper, we are specifically interested in the way in which people build and use models of their opponent to facilitate learning transfer, when engaged in situations involving an interaction with strategic considerations. These situations arise frequently such as in negotiations, auctions, strategic planning and all other domains in which theory of mind abilities (Premack & Woodruff, 1978) play a role in determining human behaviour.

In order to explore learning transfer in strategic settings, it is generally useful to study simple games as a model of more complex interactions. More specifically, when studying learning, we need a framework that allows the study of whether and how a player

takes into consideration, over time, the impact of its current and future actions on the future actions of the opponent and the future cumulative rewards. Repeated games, in which players interact repeatedly with the same opponent and have the ability to learn about the opponent's strategies and preferences (Mertens, 1990) are particularly adapted to modelling learning about the opponent.

Early literature on learning transfer in games has mostly focused on measuring the proportion of people who play normatively optimal (Nash Equilibria) or salient actions (e.g Risk Dominance) in later games, having had experience with a similar game environment previously. For instance, Ho et al. (1998) measure transfer as the proportion of players who choose the Nash Equilibrium in later p-beauty contest games, after traingn on similar games. They find there is no evidence of immediate transfer (Nash equilibrium play in the first round of the new game) but positive structural learning transfer as shown by the faster convergence to equilibrium play by experienced vs non experienced players. Camerer & Knez (2000) test learning transfer in players exposed to two games with multiple equilibria sequentially and explore the ability of players to conordinate their actions to choose a particular equilibrium in subsequent games having reached it in prior ones. They distinguish between games that are similar in a purely descriptive way, meaning similar choice labels, identity of players, format and number of action choices; and games that are similar in a strategic sense, meaning similar payoffs from combination of actions, identical equilibrium properties or significant social characteristics of payoffs such as possibility of punishment, need for fairness and cooperative vs competitive settings. They find that transfer of learning (successful coordination) occurs more readily in the presence of both descriptive and strategic similarity. If the games were only strategically similar, then the transfer was much weaker.

Juvina et al.(2014) made a similar distinction between what they deemed surface and deep similarities and find that both contribute to positive learning transfer. However,

they show that surface similarity is not necessary for deep transfer and can either aid or block this type of transfer depending on whether it leads to congruent or incongruent actions in later games. In a series of experiments using economic signalling games, Cooper & Kagel (2003, 2004, 2008) find that participants who have learned to play according to Nash Equilibrium in one game can transfer this to subsequent games, even though the actions consistent with Nash Equilibrium in later games are different. They show that this transfer is driven by the emergence of sophisticated players who are able to represent the strategic implications of their actions and reason about the consequences of changed opponent payoffs.

Most of these studies fail to offer a formal explanation of this transfer or a modelling framework that can explain the experimental observation of transfer between games and generalise it to extensive classes of games. A notable exception is the effort by Haruvy and Stahl (2012) to specify a model of learning where players learn abstract rules that they can generalise and transfer across dissimilar games, rather than action choices that can only be used within the same game. Participants played ten games, presented as 4x4 normal form (matrix payoffs). Their results suggest that subjects do transfer learning over descriptively similar but strategically dissimilar games and that this learning transfer is significant. They also showed that players learn abstract aspects of the game that are then transferred to new settings. Their rule-learning model, based on Stahl (1996), was able to capture participants dynamic behavior and show that the propensity to select particular rules is perfectly transferred across games.

In most of the experimental paradigms we saw, participants were playing other human players either one on one or as a group. In this case, players are learning about the game but so are their opponents. While the approach of matching groups of human players repeatedly to play economic games has ecological validity in recreating environments where social decision making can be studied, it complicates the detection

and modelling of strategies used by participants. It is harder to focus on an individual and how her strategies are changing and adapting to the opponent's play if we cannot experimentally control the behaviour of the opponent. It is also difficult to disentangle the process of learning about the opponent from that of learning about the game structure and payoffs. One potential solution is to replace the human opponent with a computerised agent, allowing researchers to manipulate the opponent's behaviour and observe how the human participants adapt their behaviour. For instance, Spiliopoulos (2013) made humans play constant sum games against 3 computer opponents programmed to take advantage of known patterns in human play such as imperfect randomization and heuristics use and found that human participants do adapt to the opponent they are facing. Shachat & Swarthout (2004) made human participants face computer opponents playing various mixed strategies in a zero-sum asymmetric matching pennies game. They found that the players changed their strategies towards exploiting the deviations from the MSNE, and that this exploitation was very likely if the deviation from the MSNE play was high.

One of the major limitations in studies of how humans adapt to computerised opponents is that they have mostly looked at the ability of players to detect and exploit action-based learning rules. They either used deterministic strategies that did not consider the human participant's play, as in the case of pre-determined mixed strategy algorithms, or strategies that only depended on the player's patterns of past plays. These strategies did not necessarily take into account the computer's own past plays, or the degree of strategic reasoning of the human player about the computer. Their play was not "human-like" in the sense that humans are not very good at playing specific mixed strategies with any precision of at detecting patterns from long sequences of past play due to cognitive constraints. It is therefore important to have agents that "play like humans", and one way of achieving that is to embed theses agents with human-like theory of mind abilities based on limited steps of recursive reasoning. Simon (1972) explains that humans have limited cognitive capacities and as such cannot be expected to solve computationally

₁₂₀ intractable problems such as finding Nash equilibria. Instead, they will try to "satisfice" by

₁₂₁ choosing a strategy that is adequate in a simplified model of the environment, rather than

₁₂₂ an optimal one. This concept finds its natural application in "level-k" theory, first adopted

₁₂₃ by Stahl & Wilson (1995). It posits that deviations from Nash equilibrium solutions in

₁₂₄ simple games are explained by the fact that humans have a heterogenous degree of

₁₂₅ strategic sophistication. At the bottom of the ladder, level-0 players are non-strategic and

₁₂₆ play either randomly or use a salient strategy in the game environment (Arad &

₁₂₇ Rubinstein, 2012). Level-1 players are next up the ladder of strategic sophistication and

₁₂₈ will assume all their opponents belong to the level-0 category and as such will best

₁₂₉ respond to them given this assumption. Likewise, a level-2 player will choose actions that

₁₃₀ are the best response given the belief that all opponents are exactly one level below and so

₁₃₁ on.

₁₃₂   Another potential issue in the experimental paradigms used in the learning transfer

₁₃₃ literature is the use of social dilemmas and cooperative/competitive settings that may

₁₃₄ invoke confounding factors in learning and transfer. Social dilemma and coordination

₁₃₅ games crystallise the conflict between competing with others out of self-interest or

₁₃₆ cooperating to reach a socially optimal outcome. Playing these games repeatedly

₁₃₇ implicates important social constructs such as reputation building, trust and other

₁₃₈ individual psychological attributes such as cooperativeness and inequity aversion.

₁₃₉ Coordination games also test the ability of choosing the safe self-interested choice

₁₄₀ compared to the risky cooperative choice and may depend on similar latent factors. As

₁₄₁ such, these confounding factors may impede teasing out the process of learning an

₁₄₂ opponent model and that of whether/how these models are transferred. By contrast,

₁₄₃ purely competitive settings and zero sum games in particular would be more appropriate.

₁₄₄ They do not incentivize any cooperation, since one player's gain is necessarily the

₁₄₅ opponent's loss, and are agnostic to trust mechanisms or reputation building. Also

₁₄₆ choosing games with no pure strategy Nash Equilibrium shift the focus away from learning

a normatively optimal action to that of reasoning dynamically about the opponent which

is more amenable to modelling learning. Games with unique mixed strategy Nash

equilibrium leading to nil average rewards facilitate inferring whether participants have

learned to exploit the non-random play of the opponent.

To address some of these limitations in the existing literature on learning transfer

outlined above, we propose to explore opponent modelling and its transfer with the use of

computer agents possessing human-like theory of mind abilities with limited degrees of

iterated reasoning. The agents will have a fixed strategy played stochastically and will

either be level-1 or level-2 behavioral rule, mimicking human theory of mind abilities and

the limited recursion depth they exhibit (Goodie et al., 2012). We also use a set of purely

competitive zero-sum games that incentivise reasoning about the opponent without

introducing social confounding factors. We measure transfer of learning about the

opponent strategy between games that are similar and dissimilar. The first two games we

use are identical except for action labels. In one experiment, the third game is strategically

similar to the first two but descriptively different, while in a second experiment, we

introduce a third game that is dissimilar to the first two in terms of payoff matrix and

strategic structure while retaining, like all other games, a unique mixed strategy Nash

equilibrium of random action with 0 expected reward against a Nash player.

Finally, unlike the vast majority of experimental paradigms, we make participants

play the games in an engaging, interactive and ultimately more ecological way rather than

simply provide the matrix form of the game. We use visual aids representing the player

and their computer opponent as well as the various actions that people can select built on

a highly interactive interface providing image, audio and sometimes video feedback at each

step, from choosing actions, to presenting outcomes of the rounds and players scores.

¹⁷¹ **Methods**

¹⁷²   We ran two experiments where human participants played against computer

¹⁷³ opponents a total of 3 different games.

first experiment:
¹⁷⁴

¹⁷⁵   In the first experiment, participants played 3 games sequentially against the same

¹⁷⁶ computer opponent. The computer opponent either used a level-1 or level-2 strategy. The

¹⁷⁷ three games were rock-paper-scissors, fire-water-grass and the numbers game. A typical

¹⁷⁸ rock-paper-scissors game (hereafter RPS) is a 3x3 zero sum game, with a cyclical hierarchy

¹⁷⁹ between possible actions: rock blunts scissors, paper wraps rock, and scissors cut paper. If

¹⁸⁰ one player choses an action which dominates their opponent's action, the player wins

¹⁸¹ (receives a reward of 1) and the other player loses (receives a reward of -1). Otherwise it is

¹⁸² a draw and both players receive a reward of 0. It has a unique MSNE consisting of

¹⁸³ randomly playing one of the three options each time.

¹⁸⁴   The second game is identical to Rock-Paper-Scissors in all but action labels. We call

¹⁸⁵ it Fire-Water-Grass (FWG): Fire burns grass, water extinguishes fire, and grass absorbs

¹⁸⁶ water. We are interested in exploring whether learning is transferred in a fundamentally

¹⁸⁷ similar game where the only difference is in the description of the choice actions. Finally,

¹⁸⁸ the numbers game is a generalization of rock-paper-scissors. In the variant we use, 2

¹⁸⁹ participants concurrently pick a number between 1 and 5. To win in this game, a

¹⁹⁰ participant needs to pick a number exactly 1 higher than the number chosen by the

¹⁹¹ opponent. For example, if a participant thinks their opponent will pick 3, they ought to

¹⁹² choose 4 to win the round. To make the strategies cyclical as in RPS, the game stipulates

¹⁹³ that the lowest number (1) beats the highest number (5), so if the participant thinks the

¹⁹⁴ opponent will play 5, then the winning choice is to pick 1. This game has a structure

¹⁹⁵ similar to RPS in which every action is dominated by exactly one choice. All other

¹⁹⁶ possible combination of choices that are not consecutive are considered ties. A win would

add 1 point to the score of the player, while a loss deduces one point and a tie does not affect the score. Similar to RPS, the MSNE is to play each action with equal probability in a random way.

Participants were informed they would play three different games against the same computer opponent. Each participant plays all three games consecutively and in the same order described above. Participants were told that the opponent cannot cheat and will choose its actions simultaneously without knowledge of the participant's choice. A total of 50 rounds of each game was played with the player's score displayed at the end of each game. The score was calculated as the number of wins minus the number of losses. Ties did not affect the score. In order to incentivise the participants to maximise the number of wins against the opponents, players were paid a bonus at the end of the experiment that was proportional to their final score. The overall score of the players was translated into the bonus by making each point worth \$£\$0.02. This bonus is significant as players could increase the total payoff from the experiment by up to 60% assuming they'd won all rounds against the computer opponent.

Second Experiment

In the second experiment, participants each played 3 games sequentially against both level-1 and level-2 computer opponents, rather than just one like in the first experiment. The three games were Rock-Paper-Scissors, Fire-Water-Grass, and the penalty shootout game. The first two games were identical to the ones used in the first experiment. In the final game (shootout) the participants played the role of the player shooting a football (soccer) penalty, and the AI opponent was the goalkeeper. Players had the choice between three actions, like in the first two games: Shooting the football to the left, right or centre of the goal. If the player shoots in a direction different from that of where the goalkeeper dives, they win the round and the goalkeeper loses. Otherwise, the goalkeeper catches the ball and the player loses the round. There is no possibility of ties in this game. Figure

XXXXXXXXX shows a snapshot of play in the shootout game. What makes this game different however is that there are two ways to beat the opponent in each round: if we think the opponent is going to choose "'right"' in the next round, we can win by both choosing "'left"' and "'center"'. A level-1 player (thinks that his opponent will repeat his last action) has two ways to win the next round. As such, we have programmed the level-2 computer program to choose randomly between the two possibilities that a level-1 player may choose.

Like in the first experiment, the computer opponents retained the same strategy throughout the 3 games, however the participants faced each opponent twice in each game. Specifically, each game was divided into 4 stages numbered 1 to 4, consisting of 20, 20, 10, and 10 rounds respectively for a total of 60 rounds per game. Participants start by facing one of the opponents in stage one, then face the other in stage two. This is repeated in the same order in stages 3 and 4. Which opponent they faced first was counterbalanced. All participants engage in the same three games (namely RPS, FWG and Shootout) in this exact order, and were aware that the opponent was not able to know their choices beforehand but was choosing simultaneously with the player.

In order to encourage participants to think about their next choice, a countdown timer of 3 seconds was introduced at the beginning of each round. During those 3 seconds, participants could not choose an option and had to wait for the timer to run out. A small delay that changed randomly (between 0.5 and 3 seconds) was also introduced in the time it took the AI agent to give back their response to make it seem like it is also thinking about its last actions, as a way of simulating a real human opponent. After each round, the participants were given detailed feedback about their opponent actions as well as whether they won or lost the last round. Further information about the outcome of previous rounds was also visible on the game screen below the feedback area, throughout each stage game and opponent could go back many rounds to study the history of

$_{249}$ interaction. The number of wins, losses and ties were clearly shown at the top of the

$_{250}$ screen for each game, and this scoreboard was reinitialised to zero at the onset of a new

$_{251}$ stage game. An example of the interface for the rock-paper-scissors game is provided in

$_{252}$ Figure XX.

$_{253}$ To summarise, all the games used in both experiments have a unique MSNE

$_{254}$ consisting of randomising across actions. If participants follow this strategy, or simply dont

$_{255}$ engage in learning how the opponent plays, they woud score 0 on average against both

$_{256}$ level-1 and level-2 players. Evidence of sustained wins would indicate that participants

$_{257}$ have learned to exploit patterns in the opponent play.

<div align="center">

Results

</div>

$_{258}$

First Experiment

$_{259}$

$_{260}$ In the first experiment, a total of 52 (28 female, 24 male) participants were recruited

$_{261}$ on the Prolific Academic platform. The mean age of participants was 31.2 years.

$_{262}$ Participants were paid a fixed fee of £2.5 plus a bonus dependent on their performance

$_{263}$ which averaged £1.06. The study used a 2 (computer opponent levels 1 and 2) by 3

$_{264}$ (games of rock-paper-scissors, fire-water-grass and numbers) design, with repeated

$_{265}$ measures on the second factor.

$_{266}$ Looking at the aggregate scores, the RPS game had the lowest average score across

$_{267}$ participants (M = 0.289, SD = 0.348) followed by NUMBERS (M = 0.31, SD = 0.347)

$_{268}$ and finally the FWG game had the highest average score (M = 0.454, SD = 0.354).

$_{269}$ Aggregate average scores for each game were significantly different from 0 (hypothesised

$_{270}$ value of random play) using parametric one sample t-tests (RPS: t(51) = 7.26, p-value =

$_{271}$ 2.1e-08 ; FWG: t(51) = 10.04 , p-value = 1.10e-13 ; NUMBERS: t(51) = 7.17, p-value =

$_{272}$ 2.87e-09).

$_{273}$ To analyse within and between game learning, we used a 2 (condition: level-1,

level-2) by 3 (game: RPS, FWG, NUMBERS) by 2 (block: first half, second half) repeated

measures ANOVA with the first factor varying between participants. As the number of

participants facing level-2 agents was slightly higher than those facing level-1 agents (28 vs

24), the ANOVA is unbalanced. We therefore used the "afex" (Singmann & al., 2016)

package in R, as it allows us to handle unbalanced designs and produces type III sum of

squares by default.

There was a main effect of Game ($F(2,100) = 8.54$, ges = 0.05, p = .0005), showing

that average scores varied significantly over the games. Post-hoc pairwise comparisons

showed that performance in the FWG game was significantly higher than in the RPS game

($t(100) = 3.78$, p = 0.0008), and the performance in NUMBERS was significantly lower

than FWG game ($t(100) = -3.32$, p = 0.0024). The score in RPS was not significantly

different from the score in NUMBERS however ($t(100) = 0.45$, p = 0.65). The main

effect of Block ($F(1,50) = 22.51$, p < .001, ges = 0.03) shows that the average score in

the first half of games (M = 0.29) was significantly lower than in the second half of the

games played (M = 0.40), which translates to within-game learning : as players get more

interaction with the agent, they learn how to win more often. The main effect of

Condition ($F(1,50) = 5.44$, p = .024, ges = 0.05) indicates that scores were higher against

the level-1 player (M = 0.43) than against the level-2 player (M = 0.27). This means that

it was harder for participants, on average, to learn the strategy of the more sophisticated

opponent (level-2) compared to that of the comparatively less sophisticated agent (level-1).

Finally, the analysis showed a significant interaction effects of block by game ($F(2,100) =$

6.92 , p = .002, ges = 0.02), indicating that within-game learning differed between the

games, as well as a three-way interaction between condition, game, and block ($F(2,100) =$

3.88 , p = .023, ges = 0.01).

Learning transfer.   As a measure for learning transfer, we focus on participants'

scores in the earliest rounds, excluding the very first round for which the agent is

programmed to play randomly as it has no data on prior rounds on which to build its

301 response. A group of players with no experience of the game are expected to have scores

302 not significantly different from 0. Any significantly positive group average scores would

303 therefore reflect prior learning from past experiences. In figure XXX, the average score

304 across participants by game for rounds 2-6 are plotted. Scores are also averaged across

305 levels of condition. We test the average scores for each game against a hypothesised value

306 of 0 for a non-experienced player using parametric one sample t-tests. As expected for the

307 initial RPS game, the average score is not significantly different from 0 ($t(148.85) = 1.04$,

308 $p = 0.89$). In FWG, the score is significantly higher than 0 ($t(148.85) = 4.58$, $p <$

309 0.0001). Unexpectedly, this is also the case for NUMBERS ($t(148.85) = 3.00$, $p =$

310 0.0092).

311 **Moderation of learning transfer by type of opponent and game similarity.**

312 Figure 10 shows the marginal mean scores for each game for both level-1 and level-2 facing

313 players. Graphically we can see that the pattern is dissimilar between level-1 and level-2

314 players, and we suspect transfer to be positively associated with similarity and negatively

315 with degree of sophistication of the agent. To test these hypotheses, we run statistical

316 tests on scores by game and opponent against the null hypothesis of 0 (no transfer). Table

317 XXXXXXX shows the scores on rounds 2-6 for both type of opponents in each game. We

318 also report the t-stats and p-values to test the null hypothesis that scores on these early

319 rounds are equal to 0. Rejecting the null hypothesis means that prior learning has been

320 used to beat the opponent in these early rounds. For level-1 facing players, there is

321 evidence of learning transfer to both FWG and NUMBERS. For level-2 facing players,

322 there is evidence for transfer to the similar game FWG (albeit scores are lower than for

323 level-1 player) but not to the dissimilar game of NUMBERS.

324 TABLE OF SCORES BY OPPONENT AND GAME (PG 40)

325 Our results when averaging across conditions (previous section) showed that there

326 was indeed evidence for transfer to the more dissimilar game (NUMBERS). We can see

from TABLE XXXXX that this transfer is exclusively driven by level-1 facing players, as scores of level-2 facing players are close to nil. Therefore, both groups of condition can generalise to the similar game, but only those facing the less sophisticated opponent are able to generalise to the less similar game.

Second Experiment

A total of 48 participants (21 females, 28 males, 1 unknown) used the Prolific Academic platform to participate in the experiment. The average age of players was 30.2 years, and the mean duration to complete the task was 39 minutes. Participants were incentivised using a two-tier payment mechanism: a fixed fee of £2.5 for completing the experiment plus a performance linked bonus that averaged £1.32. The participants faced two types of opponents, a level-1 and level-2 players, and played a total of 3 games (Rock-Paper-Scissors, Fire-Water-Grass, and the Penalty Shootout game), with repeated measures on the latter factor. Each game was split into 4 stages. The players faced either the level-1 or level-2 player on the first stage, and the other opponent in stage 2, for 20 rounds in each stage. Stages 3 and 4 were repetitions of the first two stages but with only 10 rounds in each stage.

The RPS game had the lowest average score per round (M = 0.194, SD = 0.345) followed by FWG (M = 0.27, SD = 0.394) and finally the Shootout game had the highest average score (M = 0.622, SD = 0.326). The higher score in shootout is expected as there are 2 out of three possible winning actions, compared to one out of three in RPS and FWG. Indeed, a player not aiming to uncover the opponent's strategy and thus choosing to play randomly should be expected to have on average score per round of 0 in both RPS and FWG, and 0.33 in the Shootout game. To make the scores more comparable, and because we are interested in player's performance that is not due to chance, we will correct all scores in the shootout game by subtracting the average score per round of a random strategy (0.33). Using parametric t-tests on adjusted scores, we reject the null hypothesis

of random play in all three games (RPS: t(49) = 6.26, p-value < 0.0001 ; FWG: t(49) =

7.25 , p-value < 0.0001 ; SHOOT: t(49) = 13.61, p-value < 0.0001).

Using the average scores obtained by participants in each game and interaction, we

explore whether learning has occurred within and between games. We perform a two

(condition: level-1 first, level-2 first) by two (opp_type: level_1 or level_2) by three

(game: RPS, FWG, SHOOT) by two (interaction: first or second) repeated measures

ANOVA with the first factor varying between participants. As the number of participants

facing level-1 agents was slightly higher than those facing level-2 agents (27 vs 23), the

ANOVA will not be balanced between the two levels of condition. We therefore used the

"afex" (Singmann & al., 2016) package in R, as it allows us to handle unbalanced designs

and produces type III sum of squares by default. The full results of this analysis are

provided in Table XXXXXXXXX.

There is evidence for a main effect of Game on scores (F(1.85,88.7) = 11.81, ges =

0.04, p < .0001), meaning that we reject the null hypothesis of scores in games being equal.

To explore these differences further, we look at post-hoc analyses for pairwise comparisons

between game scores (p-values adjusted using Holm method for multiple comparisons). We

find the performance in the games increases steadily throughout the experiment, with

FWG performance significantly higher than RPS (t(96) =2.53, p = 0.025), and

performance in the SHOOT game also significantly higher than in FWG (t(96) = 2.32, p

= 0.025). There was no main effect of either opponent type, the interaction factor(first

or second time opponent was faced) , or the condition factor (whether level-1 or level-2

opponent was faced first). This means that if we ignore potential secondary effects from

interactions, neither the type of opponent faced, nor whether it was the first or second

interaction with the opponent nor which opponent was faced first did have a main effect

on performance.

378   **Learning transfer.**   As a measure for learning transfer we will compare scores only
379 on rounds 2-6 of each game (excluding the very first round where play is necessarily
380 random). Because the number of rounds here is very limited, there should be very little
381 learning within games, and we should be better measure learning transfer if it exits. In the
382 figure below, we show the average score across participants in rounds 2-6. These scores are
383 also averaged across the levels of condition (meaning they are irrespective of which
384 opponent players faced first). For both the FWG and SHOOT games, score in the early
385 rounds of the first stage are significantly higher than 0 for both opponent types. (Level-1
386 opponent: FWG: $t(270) = 4.99$, $p < 0.0001$; SHOOT: $t(270) = 6.66$, $p < 0.0001$; Level-2
387 opponent: FWG: $t(270) = 4.40$, $p < 0.0001$; SHOOT: $t(270) = 3.21$, $p = 0.004$). Looking
388 more specifically at early scores by type of opponent faced, we confirm the result from the
389 first experiment that it is easier to transfer learning to the more dissimilar game (Shoot)
390 when facing a level 1 opponent. Indeed, while the early scores of FWG for level-1 and
391 level-2 facing players are not significantly different from each other, the score of the players
392 facing the level-1 opponent is indeed almost 0.2 point per round higher than that of
393 players facing level-2 opponents, and the difference is statistically significant ($t(144) =$
394 2.45 , $p = 0.01$). These early scores have also been adjusted to account for the fact that
395 the shootout game has higher average scores when playing randomly, and therefore this
396 difference is really due to better learning transfer and not due to chance.

<div align="center">

Discussion

</div>

397