Experimental Economics (2022) 25:1052–1078
https://doi.org/10.1007/s10683-021-09742-7

ORIGINAL PAPER

Feature-weighted categorized play across symmetric games

Marco LiCalzi[1] · **Roland Mühlenbernd**[2,3]

Received: 3 August 2020 / Revised: 4 December 2021 / Accepted: 6 December 2021 /
Published online: 28 February 2022
© The Author(s) 2022

Abstract

Experimental game theory studies the behavior of agents who face a stream of one-shot games as a form of learning. Most literature focuses on a single recurring identical game. This paper embeds single-game learning in a broader perspective, where learning can take place across similar games. We posit that agents categorize games into a few classes and tend to play the same action within a class. The agent's categories are generated by combining game features (payoffs) and individual motives. An individual categorization is experience-based, and may change over time. We demonstrate our approach by testing a robust (parameter-free) model over a large body of independent experimental evidence over 2×2 symmetric games. The model provides a very good fit across games, performing remarkably better than standard learning models.

Keywords Learning in games · Categorization · Similarity

JEL Classification C72 · C92 · D83

The replication material for the study is available at: https://doi.org/10.6084/m9.figshare.17081408.

✉ Roland Mühlenbernd
muehlenbernd@web.de

Marco LiCalzi
licalzi@unive.it

[1] Università Ca' Foscari Venezia, Venice, Italy

[2] Nicolaus Copernicus University Toruń, Toruń, Poland

[3] Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin, Germany

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

1 Introduction

Experimental game theory studies patterns of behavior for agents playing games. In particular, the dynamics and evolution of choices from players who face a stream of one-shot games is usually viewed as an instance of learning: agents refine their play as they gather more experience.

There are two major approaches to learning in games: reinforcement-based and belief-based (Feltovich, 2000). Reinforcement learning presumes that over time players tend to shift their play towards actions that have earned higher payoffs in the past (Roth & Erev, 1995; Erev & Roth, 1998). Belief-based models require a player to form beliefs about her opponents' play and choose actions with higher expected payoffs (Boylan & El-Gamal, 1993). More recently, the two approaches have been blended by experience-weighted attraction learning and its later variants (Camerer & Ho, 1999; Ho et al., 2007).

A dominant research theme has been to develop tractable models that describe (and predict) aggregate and individual play over time. This program has enjoyed a qualified success. There are simple learning models that can track "some movements in choice over time in specific game and choice contexts" (Camerer & Ho, 1999). While much work has gone into the calibration of parametric models fitting single experiments, no general theory has yet emerged.

A body of research fixates on the special case where agents face a stream of *identical* games played against different opponents. The comment by Crawford (2002, p. 11) still rings true: "almost all analyses of learning, theoretical or experimental, have concerned learning to play a single, fixed game, with past plays perfectly analogous to present ones, and past behavior taken to be directly representative of likely present behavior."

This paper puts forth a novel approach, with the goal of contributing to the study of learning over *similar* games played against different opponents. The key idea is that agents (learn to) categorize games and choose the same action for all the games placed in the same category (Mengel, 2012). Categories are based on the individual experience: two agents may have different categorizations, and these may change over time.

We demonstrate our approach with a simple model that simultaneously describes (and predicts) aggregate play over a large set of independent experiments from the literature, each based on a stream of identical 2×2 symmetric games. Our main theoretical concern is robustness and generality, rather than *ex post* calibration over a single experiment: therefore, we apply the same parameter-free version of our model over all these independent experiments simultaneously.

We test our results against suitable variants of the three major learning algorithms: experience weighted attraction, fictitious play, and reinforcement learning. The simulations based on our model match the experimental data on (overall and round-by-round) aggregate play and on the evolution of empirical behavior remarkably better than the competitors. We conclude that our simple model outperforms the major learning algorithms across a large body of evidence over identical games.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Having validated our approach over identical games, we test it against the (so far) scant experimental evidence over similar games. Our model describes the empirical evidence better than the strongest competitor. Moreover, it is the only one that can reproduce the rapid adjustments in play from an individual agent who reacts to different payoffs in a stream of similar games.

Related Literature. The literature on strategic interactions influenced by similarities is quite diverse, but there are a few unifying themes. The first theme is that players have a *coarse understanding* of their environment, usually because of bounded rationality or cognitive constraints: they cannot discriminate some elements and bundle them into categories or analogy classes. The Analogy-based Expectation Equilibrium (Jehiel, 2005) and many of its variants (Jehiel, 2020) assume that players have a coarse understanding of their opponents' moves, and thus rely on "averages" for each analogy class. Mullainathan et al. (2008) and Hagenbach and Koessler (2020) study how a coarse understanding of the message space affects persuasion and cheap talk, respectively. Gibbons et al. (2021) analyze how a shared coarse understanding may facilitate or impede cooperation. Grimm and Mengel (2012) provides experimental evidence on how complexity and feedback affect whether agents's coarse understanding lead them to bundle their beliefs about opponents' moves or their own choice of actions across similar games.

A second theme is *similarity-based reasoning*, used to draw inferences and decide a course of action across different situations. The seminal contribution is case-based decision theory (Gilboa & Schmeidler, 1995, 2001). Steiner and Stewart (2008) demonstrates how similarity-based reasoning can spread "contagiously" a mode of play across increasingly different global games. Similarity-based reasoning is also invoked to explain experimental evidence on the emergence of "conventions" (Rankin et al., 2000; Van Huyck and Stahl 2018). These contributions are consistent with agents learning *within* the bounds of their coarse understanding or their similarity judgments, but do not tackle the issue of *learning categories or similarities* for games.

This goal characterizes the third theme. Mengel (2012) applies evolutionary learning to analyze how agents' categorizations over a set of games change and possibly stabilize. Heller and Winter (2016) recast such categorizations as part of the agents' strategic options. These works assume a quite small set of competing categorizations. Recent works expand agents' search over wider sets of categorizations: LiCalzi and Mühlenbernd (2019) study the evolution of binary interval partitions over a space of games that can be mapped to a one-dimensional interval; Daskalova and Vriend (2021) model the categorization of the player's own actions in a one-shot game under reinforcement learning, and test its fit with the experimental evidence. Our paper advances a simple mechanism to generate (and adapt) individual probability distributions over alternative categorizations, validated by a competitive test over independent experimental evidence.

The fourth theme collects a few ambitious attempts to consider *learning to reason across similar games*. Samuelson (2001) uses finite automata to model how agents may select among their stock of models what works better across different games. Lensberg and Schenk-Hoppé (2021) recast this problem as the search for a solution concept, using genetic programming to sort out individual selections and develop

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

an aggregate solution concept. Haruvy and Stahl (2012) provide some experimental evidence that more sophisticated players tend to move from non-belief-based to belief-based rules of play.

2 Categorization of similar games

Consider the payoff matrix for the row player in a symmetric 2×2 game:

	H	L
H	a	b
L	c	d

We assume that the payoffs are positive and not all identical. The payoffs are arranged to favor the first row, in the sense that $a > d$ (or, when $a = d$, the sum $a + b \geq c + d$.) Actions are accordingly labeled H (high) and L (Low). Let $\mathcal{G}$ denote the set of symmetric games that fit these assumptions.

If one pays attention only to a player's best replies (and ignores inconsequential ties), there are four classes of games in $\mathcal{G}$; see Rapoport and Guyer (1966) for a more refined classification. Broadly speaking, the four classes can be distinguished as follows: the first and second class have a unique equilibrium (in dominant strategies), respectively at (L, L) and at (H, H); the third and fourth class have three equilibria, of which two are in pure strategies, respectively on the main and on the secondary diagonal. We find it easier, however, to identify the four classes by their most representative game.

When $c > a$ and $d > b$, L is the dominant strategy and the game is a Prisoners' Dilemma (PD), where H and L correspond to cooperation and defection; we denote this first class by PD. When $a > c$ and $b > d$, H is the dominant strategy: the most representative game is the Prisoners' Delight (DE) where H is the cooperative choice; we denote this second class by DE. When $a > c$ and $d > b$, we have a coordination game with two equilibria in pure strategies at (H, H) and at (L, L): the most representative example is the Stag Hunt (SH) where H is the payoff dominant choice and L is the safe choice; we denote this third class by SH. When $c > a$ and $b > d$, we have an anti-coordination game with two equilibria in pure strategies at (H, L) and at (L, H): the most representative example is the Chicken Game (CG); we denote this fourth class by CG. The following table summarizes the four classes.

	$d > b$	$d < b$
$a > c$	SH	DE
$a < c$	PD	CG

We use these four sets of games as similarity classes: two games from the same class are similar (because they have the same best replies), but games from different classes are dissimilar. For instance, PD usually typifies obstacles to cooperation while SH exemplifies issues of coordination.

A player facing a game from $\mathcal{G}$ has to decide whether to play H or L. We model this decision as if the player is sorting the games into two main categories: the

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

H-group and the L-group. If the player assigns a game to the H-group, he chooses H; otherwise, he chooses L.

We propose a model of (stochastic) categorization by which a game is attributed to either group. Our model has three distinctive elements: (1) the assignment depends both on the objective features of the game and on the subjective attitudes of the player; (2) the assignment rule is stochastic, so that a game may generally be assigned to either class with some (varying) probability; (3) the assignment may change over time, because the player reviews his attitudes in view of the past history of play.

The model sorts games from $\mathcal{G}$ into an H-group and a L-group. Within a similarity class, this sorting may be viewed as a reduced form for the process of choosing to play different strategies in similar games. When the model is applied over games from different similarity classes, the model allows for some transfer of experience across distinct classes of similarity (Knez & Camerer, 2000).

Our work is inspired by an impressive meta-study of 96 laboratory experiments concerned only with the PD class, where Mengel (2018) identifies three key descriptors for predicting behavior in PDs: temptation, risk and efficiency. These three features take values in [0, 1] and prove very useful to predict the dynamics of cooperation rates when agents face a stream of identical PD games. Mengel's three descriptors suggest that actual players pay attention to three elements: the best reply when the opponent plays H, the best reply when she plays L, and the two payoffs a and d on the main diagonal.

Best replies are standard fare in game theory, whereas the importance of the common payoffs on the main diagonal is usually underplayed. A theoretical argument supporting the relevance of symmetric (non-equilibrium) outcomes in symmetric games, suggested by Rapoport (1966) and taken up by Davis (1977), has been popularized under the name of *superrationality* by Hofstadter (1983).

We keep Mengel's terminology but extend her definitions from the PD class to the larger set $\mathcal{G}$. All the descriptors are positively or negatively signed, and normalized either to [0, 1] or to $[-1, 0]$ after dividing them by the greatest payoff in the game $M = \max\{a, b, c, d\}$. We describe a game from $\mathcal{G}$ using three features:

Temptation: $T = \frac{c-a}{M}$
is the normalized gain/loss when playing L (instead of H) against H;
Risk: $R = \frac{d-b}{M}$
is the normalized gain/loss when playing L (instead of H) against L;
Efficiency: $E = \frac{a-d}{M}$
is the normalized gain from coordinating on H versus L.

Note that E is always positive by construction, whereas T and R can take either positive or negative values. The following table summarizes their signs for the four similarity classes in $\mathcal{G}$.

	$R > 0$	$R < 0$
$T < 0$	SH	DE
$T > 0$	PD	CG

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

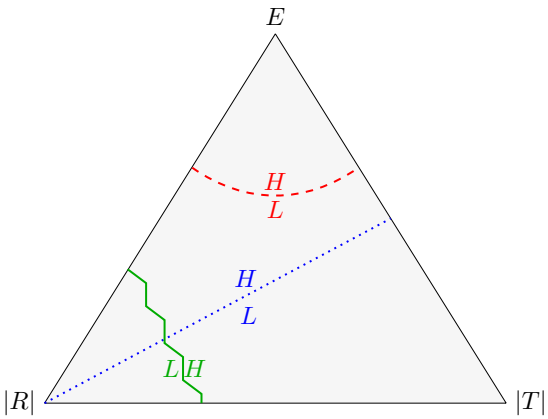


Fig. 1 The red and blue lines describe two typical categorizations for the PD-simplex: the red agent cooperates when efficiency (E) is high; the blue agent ignores risk (R) and cooperates if efficiency is greater than temptation ($E > T$). The green line gives a typical categorization for the SH-simplex: the agent plays safe when risk (R) is high

The three features (T, R, E) are *extrinsic*: they describe characteristics of the game that are independent of context and agents' experience. The sign of each feature is constant within a class; for instance, we have $T > 0$ and $R > 0$ in any game from PD, whereas $T < 0$ and $R > 0$ in any game from SH. We can use the three extrinsic features (from now on, *e-features*) to locate any PD in a 3-dimensional PD-simplex, by mapping the vector (T, R, E) to the point

$$\left(\frac{|T|}{|T| + |R| + |E|}, \frac{|R|}{|T| + |R| + |E|}, \frac{|E|}{|T| + |R| + |E|} \right)$$

Under this mapping, two games from the same class with proportional e-features (T, R, E) correspond to the same point. Likewise, for each class C =PD, SH, CG, DE, we can associate a game from C to a point in the C-simplex. The four C-simplices are formally similar, but the original T and R descriptors carry different signs for different classes of games.

Within each similarity class, higher proximity of the e-features (T, R, E) operationalizes higher similarity of the corresponding games. A categorization ascribes games to a preferred action (H or L). Figure 1 portrays three examples, where the boundary between two categories is typically a curve.

3 Feature-weighted categorized (FWC) play

An agent playing a one-shot game has to decide whether to choose H or L. Our goal is to describe the experimental results from others' studies and predict the probability of choice for either action (Erev and Roth 1998).

We model how an agent categorizes games and ascribes them to a preferred action. When an agent evaluates the strength of the arguments favoring H versus L, he weighs the e-features of the game by the strength of three individual motivations:

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Fear: Avoid the lower payoff if the opponent plays L;
Greed: Attain the higher payoff if the opponent plays H;
Harmony: Coordinate on the best symmetric payoff (on the main diagonal).

We denote the relative strength of each motivation by three weights $f, g, h \geq 0$, with $f + g + h = 1$. The relative strengths (f, g, h) of the motivations are *intrinsic* features (from now on, *i-features*) of the agent, because they pertain to his individual attitude and experience.

The *e*-features of a game and the *i*-features of an agent interact in ordered pairs: risk with fear, temptation with greed, and efficiency with harmony. We model their pairwise complementarity using a simple product operator. Each of the three interactions generates an agent's *disposition* towards playing H or L in $[0, 1]$:

$d_R = f \cdot R$ is the disposition to avoid the lower payoff when the opponent plays L;
$d_T = g \cdot T$ is the disposition to achieve the higher payoff when the opponent plays H;
$d_E = h \cdot E$ is the disposition to coordinate on the best symmetric payoff.

The disposition d_E is positive by construction and attracts the agent towards playing H (or restrain the agent from playing L). The dispositions d_R and d_T may instead be positive or negative. A positive disposition d_R or d_T attracts the agent towards playing L (or restrain the agent from playing H); a negative disposition d_R or d_T restrain the agent from playing L (or attracts the agent towards playing H). We denote the positive part of a number x by $(x)^+ = \max(0, x)$, the negative part by $(x)^- = \max(0, -x)$, and its absolute value by $|x|$.

The agent's inclination to associate a game with the choice of L (or H) combines the signed effects of the three dispositions. This can be succinctly expressed by assuming that the probability $P_i(H)$ that an agent i with i-features $(f, g, h)_i$ chooses H in a game G with e-features $(T, R, E)_G$ correlates with a positive disposition towards efficiency and with negative dispositions towards temptation and risk:

$$P_i(H) = \frac{d_E + (d_T)^- + (d_R)^-}{d_E + |d_T| + |d_R|} \tag{1}$$

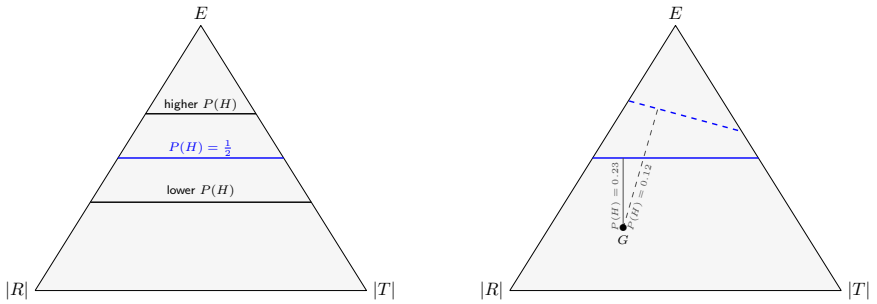
Consequently, the probability that L is chosen (or, equivalently, H is *not* chosen) correlates positively with positive dispositions towards temptation and risk

$$P_i(L) = 1 - P_i(H) = \frac{(d_T)^+ + (d_R)^+}{d_E + |d_T| + |d_R|}$$

This formulation unifies the directions of influence for each class.

Given a fixed set (f, g, h) of i-features, this (stochastic) choice rule uses the i-features of a game to select an action. When $P_i(H) = 1/2$, an agent is equally likely to attribute action H or L to a game: the equation $P_i(H) = 1/2$ defines an *ideal boundary* (IB) in the simplex that separates the region where H is more likely from the region where L is more likely. Under our assumptions, the ideal

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press



(a) The ideal boundary of the PD class **(b)** Moving the ideal boundary of the PD class

Fig. 2 **(a)** The blue line depicts the ideal boundary (IB) in the PD-class for weights $(f, g, h) = (1/3, 1/3, 1/3)$, where $P(H) = \frac{1}{2}$. **(b)** If the weights change to (0.5, 0.3, 0.2), the IB shifts from solid blue to dashed blue. The distance from a game to IB is inversely proportional to $P(H)$: f.i., if G has payoffs $a = 7, b = 0, c = 12, d = 4$, then $P(H) \approx 0.23$ for the solid IB and $P(H) \approx 0.12$ for the dashed IB

boundary IB is linear. As we move away from IB, the probability $\max(P_i(H), P_i(L))$ of the modal choice increases; see Fig. 2a.

The choice rule of the model allows for special cases, known in the literature. For example, an agent with no harmony motivation ($h = 0$) has a zero disposition towards efficiency ($d_E = 0$) and therefore always defects in a PD; this behavior is frequently observed in laboratory experiments (Andreoni & Miller, 1993). Moreover, if this agent is equally motivated by fear and greed ($f = g = 0.5$), she would choose the risk-dominant option in any SH. As a different example, an agent with no greed motivation ($g = 0$) would be unable to distinguish PD from SH if the risk and efficiency features take the same values (Devetag & Warglien, 2008).

The choice rule (1) describes the probability that an agent facing a game G at a given period chooses H or L. However, the strengths of his motivations (and hence the agent's i-features) change over time in accordance with his past experience. Therefore, the probabilities may be updated after each play.

Our update rule distinguishes three (qualitative) situations. Two are valid across all games, while the third applies to two classes of games: each update rule concerns one of three motivations. First, after a recent history of opponents playing L, the agent becomes more concerned about avoiding the lower payoff against L: this increases the weight f for fear. Second, after a recent history of opponents playing H, the agent is more tempted to achieve the higher payoff against H: this increases the weight g for greed. Third, after the opponent's last choice of H, the agent feels more motivated to contribute towards the efficient symmetric profile (H, H) if this yields the best possible outcome (as it is for SH and DE): this increases the weight h for harmony.

We assume that the update rule is based on the last observed opponent's action s_j^t. At time t, the current strengths $(f, g, h)_i$ for i are updated using the rule

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

$$(f', g', h')_i = \begin{cases} \left(f + \gamma_i,\ g,\ h\right)_i & \text{if the last opponent's choice is } (L_j^t) \\ \left(f,\ g + \gamma_i,\ h\right)_i & \text{if the last opponent's choice is } (H_j^t) \\ \left(f,\ g,\ h + \gamma_i\right)_i & \text{if the last opponent's choice is } (H_j^t) \text{ and } (H, H) \text{ gives the highest possible payoff} \end{cases} \tag{2}$$

and afterwards they are divided by $(1 + \gamma_i)$ to renormalize their sum to 1. After an update, the IB shifts and changes the implicit categorization used by the agent; see Fig. 2b.

The update rule reacts to the opponent's play by increasing the weight of the concern for best replying to the last observed move. Moreover, when the opponent's last move H can be plausibly interpreted as an attempt to coordinate on the highest payoffs, the update rule also increases the weight for harmony. The magnitude γ_i of the update represents the *learning rate* (or the reaction speed) of agent i. Because games closer to the IB are more informative, we let this rate decrease in the probability of the modal choice and set $\gamma_i = 1 - \max(P_i(H), P_i(L))$.

4 FWC play over the PD and SH classes

The largest number of experiments on learning in symmetric 2×2 games concerns games from the PD-class and from the SH-class. This section covers a comprehensive examination of the FWC dynamics that focuses on these two game classes. We investigate (i) experiments of identical games, (ii) experiments of similar games that belong to the same game class, and (iii) an experiment that involves PD and SH games. In the subsequent section, we test our algorithm against the rather sparse empirical evidence of games from other classes, including chicken games and some 3×3 games.

4.1 Identical games

We test the descriptive and predictive power of FWC dynamics against laboratory evidence collected by others. We selected all the 26 PD treatments in Mengel (2018) and other 19 SH treatments from the experimental literature that fulfill three criteria: (a) the stage game has positive payoffs; (b) the stage game is played over multiple rounds; (c) players are paired using the random matching protocol. See Table SM.1 for a full list of the selected games. (The prefix SM refers to items listed as Supplementary Material in the Online Appendix, for which a link is provided at the end of the paper.)

After mapping each game to its three e-features (T, R, E), Fig. 3 shows the 26 PD games (left panel) and the 19 SH games (right panel) projected on the (T, R, E)-simplex.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

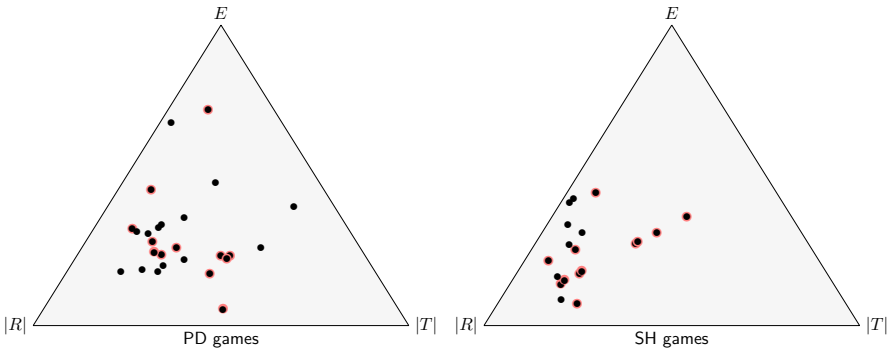


Fig. 3 We test 26 PD games (left panel) and 19 SH games (right panel) for overall average rates of *H*-play. Additionally, we test 12 of the PD games and 12 of the SH games (red rings) for round-by-round average rates of *H*-play. Experimental evidence for these 45 games is collected from 16 different studies. See Table SM.1 for a full list of games and sources

We evaluate the performance of the FWC model by means of a comparative test against three major learning algorithms: experience weighted attraction (EWA), fictitious play (FP), and reinforcement learning (RL). As a benchmark, we also simulate the outcome for two fixed strategies: the coin flip (CF) that in every round plays each action with equal probability, and the Nash equilibrium (NE) that plays L in any game from the PD class and the equilibrium in mixed strategies in any game from the SH class.

We implemented plain-vanilla (parameter-free) versions for each learning algorithm where agents' choices are stochastic, making no attempt to calibrate a model to the empirical evidence and trying our best to avoid biases. For instance, each agent was initialized with the same equal weights ($f = g = h = 1/3$) in FWC. A full description of the three algorithms (FP, RL, EWA) can be found in Appendix A.1. (The prefix A refers to items in the Online Appendix, for which the link is provided at the end of the paper.)

We ran computer simulations matching the original experimental settings: in particular, we replicated payoffs, number of participants, and number of repetitions. For each algorithm, we computed the *overall mean squared error* (MSE) between simulated data and experimental data, defined by

$$Q_1 = \frac{1}{n}\sum_{k=1}^{n}\left(E_k - S_k\right)^2$$

where n is the number of experiments, E_k is the overall average rate of H-play (H-rate, for short) for the game in the k-th experiment, and S_k is the overall average H-rate for the same game after 1000 simulations (Selten, 1998; Feltovich, 2000; Chmura et al., 2012).

Figure 4 summarizes the results concerning $Q1$. (The actual values are listed as Table 1.) The two rightmost columns in each panel correspond to the CF and NE benchmarks and can be used to gauge the improvement brought by the models

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Table 1 Values of the mean square distances Q_1 and Q_2 for FWC, EWA, FP, RL, CF and NE.

	FWC	EWA	FP	RL	CF	NE
Q_1 (26 PD)	0.029	0.033	0.031	0.044	0.088	0.073
Q_1 (19 SH)	0.043	0.048	0.081	0.074	0.087	0.238
Q_1 (all 45)	0.035	0.04	0.052	0.057	0.087	0.143
Q_2 (12 PD)	0.021	0.026	0.028	0.04	0.102	0.063
Q_2 (12 SH)	0.023	0.077	0.099	0.116	0.117	0.249
Q_2 (all 24)	0.022	0.055	0.063	0.078	0.11	0.156

Each data point is based on 1000 simulation runs. These data are used to obtain Figs. 4, 6

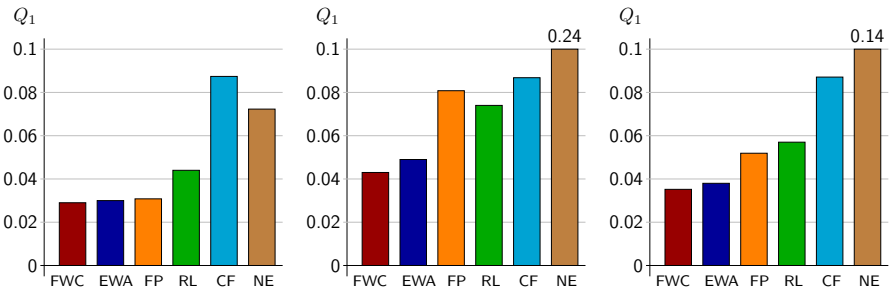


Fig. 4 Comparative evaluation between experimental results and simulated data over six algorithms, using the mean square deviation Q_1 for the overall average rate of H-play. The six algorithms are: *feature weighted categorization* (FWC), *experience weighted attraction* (EWA), *fictitious play* (FP), *reinforcement learning* (RL), *coin flip* (CF), and *Nash equilibrium* (NE). The three panels show the values for Q_1 computed over: (a) 26 PD games, (b) 19 SH games, and (c) all 45 games together (black dots in Fig. 3). When the value is off scale, we add a number on top of the bar

in the four leftmost columns (FWC, EWA, FP, RL). Panel 4a shows the Q_1 values over 26 PD games, Panel 4b shows the Q_1 values over 19 SH games: FWC outperforms the other models for both classes, scoring slightly better than the second best algorithm (EWA). Panel 4(c) shows the Q_1 values over the whole set of 45 games.

This first batch of results concerns the overall rates of H-play, averaged across agents and over all rounds of play. The time average, however, may disguise the dynamics of behavior. For instance, if the H-rate in the experiment decreases linearly from 1 to 0 and the simulated H-rate increases linearly from 0 to 1, the dynamics would diverge but the MSE as measured by Q_1 would be nil. A more exacting measure of distance between experimental data and simulated data is the *round-by-round mean squared error* Q_2, defined by

$$Q_2 = \frac{1}{n}\sum_{k=1}^{n}\left(\frac{1}{T_k}\sum_{t=1}^{T_k}\left(E_k^t - S_k^t\right)^2\right)$$

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

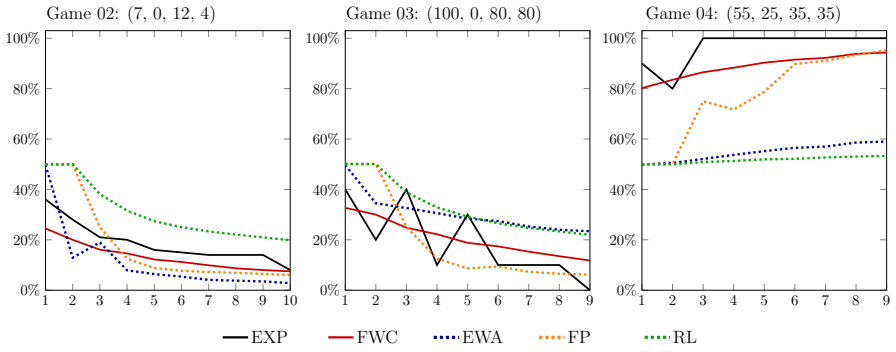


Fig. 5 Round-by-round average *H*-rates over three identical games: Game 02 (left) is PD, Game 03 (center) is SH and participants learn to play *L*, Game 04 (right) is SH and participants learn to play *H*. We depict the experimental results (solid black line) and the simulated *H*-rates averaged over 1000 simulation runs for FWC (red line), EWA (blue line), FP (orange line) and RL (green line)

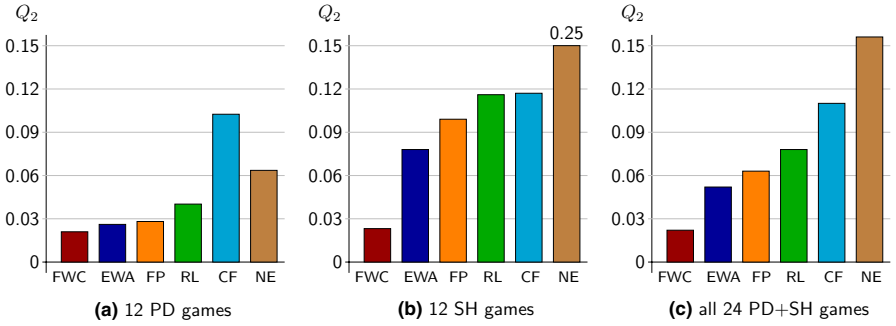


Fig. 6 Comparative evaluation of the mean squared deviation between experimental results and simulated data using Q_2 for the round-by-round average rate of *H*-play, over the same six algorithms of Fig. 4. The three panels show the values for Q_2 computer over: **(a)** 12 PD games, **(b)** 12 SH games, and **(c)** all 24 games together (red rings in Fig. 3). When the value is off scale, we add a number on top of the bar

where n is the number of experiments considered, E_k^t and S_k^t are respectively the average *H*-rate in the t-th round of the k-th game from the experiment and from 1000 simulations, and T_k is the number of experimental rounds for the k-th game.

We ran a second comparative test using Q_2 over a smaller dataset, because the round-by-round *H*-rates are available only for 12 PD games and 12 SH games from our original selection of 45. These games are marked with red rings in Fig. 3. Table SM.2 provides the full list of values for the round-by-round average cooperation rates in each experiment we tested. Figure 5 provides an exemplary selection of the round-by-round average *H*-rates from three experiments. The evidence from 24 experiments is depicted in Fig. SM.1 (12 PD games) and Figure SM.2 (12 SH games).

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Figure 6 summarizes the results for Q_2, confirming the Q_1 measures; see Table 1 for the actual values. Panel Fig. 6a shows the Q_2 values over 12 PD games: FWC performs clearly better than the other algorithms. Panel 6b shows the Q_2 values over 12 SH games: the FWC algorithm clearly outperforms all other algorithms, delivering a Q_2 less than 30% of the second best (EWA). Panel Fig. 6c shows the Q_2 values over the whole set of 24 games.

Finally, we compare the ability of FWC and the other four learning algorithms to match two stylized facts concerning the evolution of the H-rate in a PD or in a SH, respectively; see Flache and Macy (2006) for an insightful discussion. The initial H-rate for both PD and SH is positive. But in a PD this rate decreases over time as if the players move towards the only equilibrium. In a SH, instead, this rate approaches one or zero as if the players coordinate on one of the two equilibria in pure strategies. We summarize these stylized facts with the following two propositions, numbered for later reference.

($\mathbf{A}_1$) *The H-rate in a PD declines over time.*
($\mathbf{A}_2$) *The H-rate in a SH approaches zero or one over time.*

We use the following consistent approach to evaluate the ability of the algorithms to match the content of each proposition. We choose an appropriate measure of fitness. We let $v_0 \geq 0$ be the higher fitness value achieved by either of the two benchmarks (NE or CF) and we let v_1 be the fitness value corresponding to a perfect match with the empirical data. Given the fitness value v_α achieved by an algorithm α, we assess it by its (relative) rate of improvement:

$$\frac{v_\alpha - v_0}{v_1 - v_0}$$

Conventionally, we say that an algorithm is *consistent with a proposition* A_i if its percentage rate of improvement is at least 0.5: that is, if it yields data that improve over the best benchmark by at least 50%. In the next section, we use the same approach to test three other propositions summarizing the experimental evidence over similar games.

Concerning A_1 and A_2, we evaluate the fitness using $1 - Q_2$. For example, v_0 is the higher value for $1 - Q_2$ achieved by either benchmark (NE or CF) and $v_1 = 1$ is the value for $1 - Q_2$ corresponding to a perfect match with the empirical data ($Q_2 = 0$). Specifically, the test for A_1 (and, respectively, for A_2) considers the 12 PD (SH) experiments for which the round-by-round average H-rates are available; see Table SM.2.

The first two rows in Fig. 7 summarize our results for A_1 and A_2, respectively. Each column shows graphically the percentage rate of improvement for the four algorithms: feature-weighted categorized play (FWC), experience weighted attraction (EWA), fictitious play (FP), and reinforcement learning (RL). We find that FWC, EWA and FP are consistent with A_1 (first row); but only FWC is consistent with A_2 (second row).

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

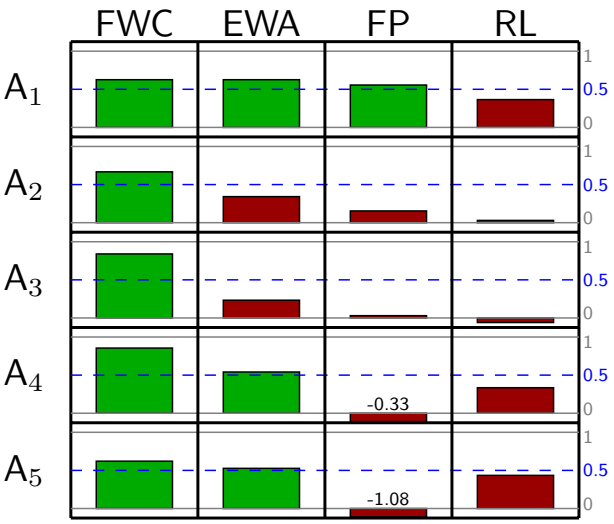


Fig. 7 Visual summary for five propositions (rows) across four algorithms (columns). Each bar depicts the relative rate of improvement of an algorithm over the best benchmark (either CF or NE) in matching the empirical evidence for a proposition. An algorithm is consistent with a proposition if its rate of improvement is above 0.5 (green), and is not otherwise (red). See Table 3 for the actual values

Overall, these comparative evaluations over a reasonably rich dataset of experiments strongly suggest that FWC is an effective approach to describe or predict the average *H*-rates for games in the PD or SH class, in a setting where agents face the same *identical* game. We view this positive outcome of our tests in a single-game setting as a necessary check of soundness. From a theoretical viewpoint, we believe that the real strength of FWC (compared with other models in the literature) lies in its ability to model the richer and more interesting case where agents face a stream of *similar* games. The next section tests the performance of FWC in describing learning over *similar* games, using experimental evidence independently collected by others; see Table SM.3 for the full list.

4.2 Similar games

The experimental evidence over participants playing a stream of similar one-shot games is sparse, and we look forward to more research in this area. We use only data available in the literature and we apply the same methodology used for A_1 and A_2, comparing the performance of the four learning algorithms with respect to three distinct propositions concerning play over *similar* games in a multi-game environment.

Schmidt et al. (2001) report an experiment where participants played 6 different PD games in a sequence of 24 rounds. The solid black line in Fig. 8 depicts the round-by-round *H*-rate observed over 4 experimental sessions with 8 participants each; see SQ_1 of Table SM.4. Its jagged shape makes it apparent that a sizable

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

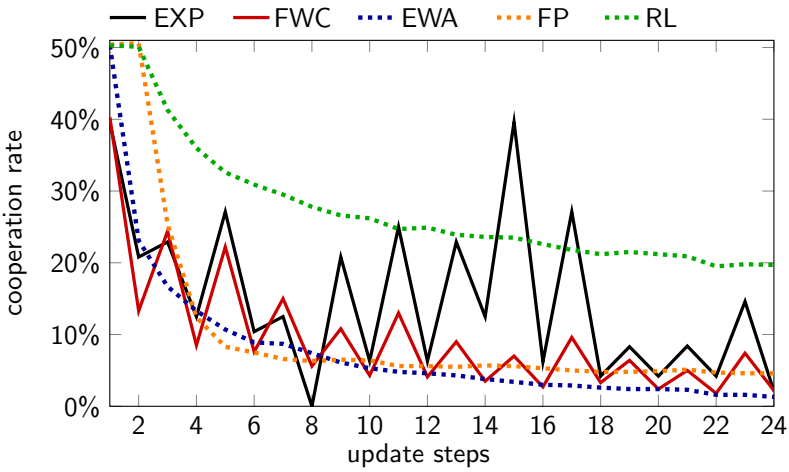


Fig. 8 Dynamics of the *H*-rates over different PD games. Experimental data (solid black line) exhibit rapid adjustments. Only FWC (solid red line) matches this pattern

number of participants make rapid adjustments to their play according to the payoffs of the game they face. We summarize this effect in the following proposition.

($\mathbf{A}_3$) *The H-rate may switch rapidly across similar games.*

The FWC algorithm is sensitive to the *e*-features of the current game and thus can yield rapid adjustments. The solid red line in Fig. 8 depicts the round-by-round *H*-rate generated by FWC for the same sequence of games as in Schmidt et al. (2001). Comparing it against the black solid line shows that the *H*-rate moves in the same direction, with jumps of similar magnitude. None of the other three learning algorithms can deliver rapid adjustments.

We test for A_3 following the same approach described above, using correlations over changes in *H*-rates (see Appendix A.2 for details about the definition and Table 2 for the correlation data). As shown in the third row of Figure 7, the test confirms that only FWC is consistent with A_3.

Rankin et al. (2000) report an experiment where six groups of eight participants played a sequence of 75 similar stag hunt (SH) games under random pairwise matching. In each of the games, the payoff matrix for the row player is the following:

	H	*L*
H	1	0
L	*x*	*x*

with $0 < x < 1$. These games differ only in the value of x: they can be viewed as elements of a one-dimensional space of SH games, fully ordered by x. Note that each game has $T = E$.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

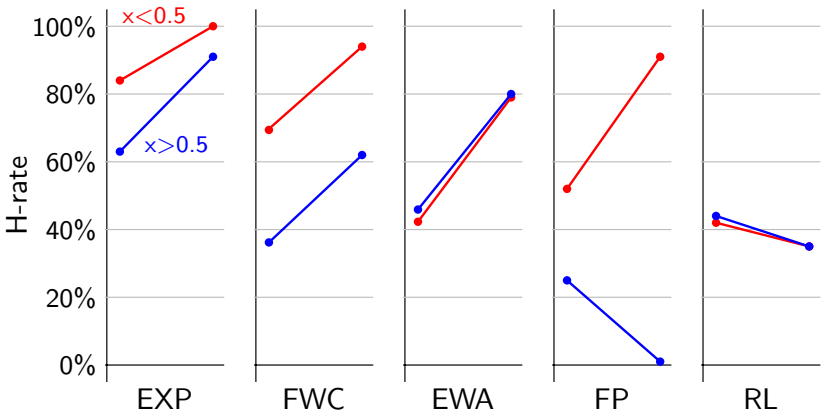


Fig. 9 Change in *H*-rates from the first 10 to the last 10 rounds for low-x SH games (x<0.5, red lines) and high-x SH games (x>0.5, blue lines) averaged over all experimental sessions and over 1000 simulations for each algorithm

Each experimental session exhibits a similar pattern. Players initially play *H* more often in games with low *x*-values ($x < 0.5$) than in games with high *x*-values ($x > 0.5$). Over time, this difference disappears; in the last rounds, participants tend to play *H* for (almost) every game as if they had developed a convention to play *H* throughout. (Participants occasionally seem to learn the opposite convention and focus on *L*; see Van Huyck and Stahl (2018)).

The first column in Fig. 9 depicts the variation in *H*-rates for low-*x* and high-*x* games between the initial ten rounds and the final ten rounds, averaged over all experimental runs; see SQ_2 in Table SM.4. We summarize this effect in the following proposition.

($\mathbf{A}_4$) *Play over distant (but similar) games is initially different but converges to a convention.*

The four columns in Fig. 9 after the first show the data generated by 1000 simulations for the same experimental setup. Each column represents one of the four learning algorithms. Only FWC matches both the initial gap in *H*-rates and the convergence to the *H*-convention. EWA delivers the *H*-convention, but fails to exhibit an initial gap.

We test for A_4, using the average mean square distance for the ratio of the differences in average *H*-rates; see Appendix A.2 for details and Table 2 for the performance values. As shown in the fourth row of Fig. 7, the test confirms that both FWC and EWA are consistent with A_4.

4.3 Dissimilar games

Finally, we consider a situation where agents first encounter identical (or similar) games from the same class, and immediately after play identical (or similar)

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Table 2 Performance of FWC, EWA, FP, RL, CF and NE for the experiments across similar games.

	measure	FWC	EWA	FP	RL	CF	NE
PD sequence	r	0.727	0.233	0.017	−0.11	0	0
SH sequence	Q_4	0.329	0.332	0.971	0.475	0.721	0.756
SH-PD sequence (no reset)	Q_2	0.143	0.231	0.244	0.152	0.09	0.158
SH-PD sequence (reset)	Q_2	0.029	0.078	0.145	0.086	0.09	0.158

We use the fitness measures defined in A.2. Each datapoint is based on 1,000 simulation runs

Table 3 Values of the relative scores for propositions A_1 to A_5 over the four algorithms FWC, EWA, FP, and RL.

	FWC	EWA	FP	RL	CF	NE
A_1	**0.667**	0.492	**0.556**	0.365	–	0
A_2	**0.803**	0.342	0.154	0.009	0	–
A_3	**0.727**	0.233	0.017	−0.11	–	0
A_4	**0.542**	**0.54**	−0.347	0.341	0	–
A_5	**0.673**	0.132	−0.611	0.045	0	–

The lowest benchmark (either CF or NE) is marked by 0. Scores above the benchmark level of 0.5 are in boldface. Each data point is based on 1,000 simulation runs. These data are used to obtain Fig. 7

games from a different class. This setup has been used to explore the possibility of *precedent transfer*, by which the mode of play developed over the first class may affect the mode of play adopted over the second class. For example, after playing coordination games and developing a convention that focuses on the payoff-dominant outcomes, agents facing PD games exhibit much higher H-rates (Knez & Camerer, 2000; Rusch & Luetge, 2016). When some transfer occurs, the literature has often labelled it as a *spillover effect* (Ahn et al., 2001; Peysakhovich & Rand, 2016).

The most detailed experimental study that is currently available and fits our setup is Duffy and Fehr (2018). They consider indefinite play of identical PD or SH games under the random matching protocol and report H-rates over four treatments, each involving one PD game and one SH game. The payoffs for all games involved are constant, except for the "temptation" payoff denoted as c in this paper. In a nutshell, they report little evidence of transfer precedent: in particular, they find no evidence for efficient play in SH to carry over to PD, and only two instances where inefficient play transfers from PD to SH.

This evidence is generally in contrast with the related literature on learning spillovers and on precedents as selection devices; see the review in Duffy and Fehr (2018, Section 2). Commenting on this discrepancy, the conclusions in Duffy and Fehr (2018) hypothesize that "One possible explanation for the lack of precedent transfer in our experiment is that our experimental design […] may have triggered an experimenter demand effect wherein subjects felt compelled to respond to the change by playing differently." The occurrence of an experimenter

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

demand effect would (at least, partially) mark a break when a different game is introduced.

While more research is needed to settle this issue, we put at use the detailed dataset in Duffy and Fehr (2018) and compare their experimental outcomes against the simulated outcomes from our four learning algorithms. We consider two extreme cases: if there is no break when a new game is introduced, the algorithms continue using the values from the last iterations; if there is a break, we reset these values to the initial condition. The outcomes with and without reset for the first sequence in Duffy and Fehr (2018) are displayed in Fig. 10. (Analogous displays for all the four sequences are in Fig. SM.3.)

The two panels report the simulated outcomes for round-by-round average *H*-rates:

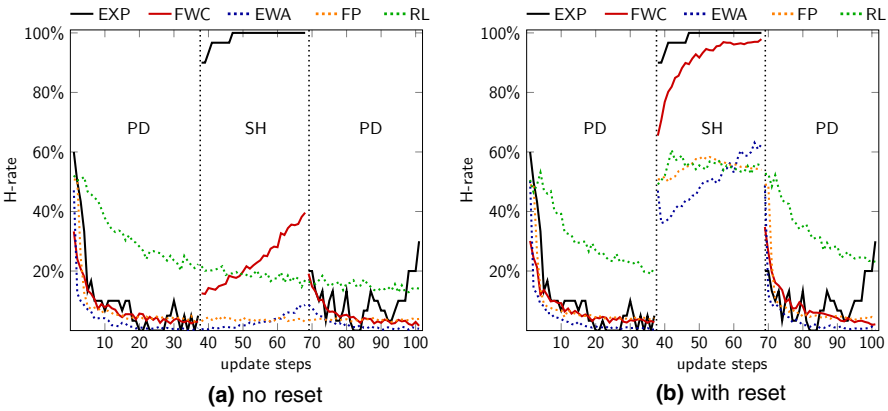


Fig. 10 Round-by-round average H-rates over the first sequence of PD and SH games from Duffy and Fehr (2018). We depict the experimental results (solid black line) and the simulated *H*-rates averaged over 1000 simulation runs for FWC (red line), EWA (blue line), FP (orange line) and RL (green line). On the left, algorithms are never reset; on the right, they are reset upon introducing a different game

on the left, when algorithms are never reset; on the right, when algorithms are reset upon introducing a game from a different class. The experimental data, depicted with a solid black line, are the same in both panels: the *H* rates over different classes are markedly different, and show no spillover effect. We summarize this effect in the following proposition.

($\mathbf{A}_5$) *The setup in* Duffy and Fehr (2018) *inhibits precedent transfer.*

Figure 11 collects Q_2 values for the round-by-round average *H*-rates of the all four algorithms, with and without reset (see Table 2 for exact values). The experimental outcomes cover a total of 418 rounds, from four treatments spread over about 100 rounds each. The simulated data for each algorithm are based on 1000 runs for either condition (no reset and with reset). Comparing simulated outcomes with experimental outcomes, all four algorithms perform better with reset. FWC performs best in either setting.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

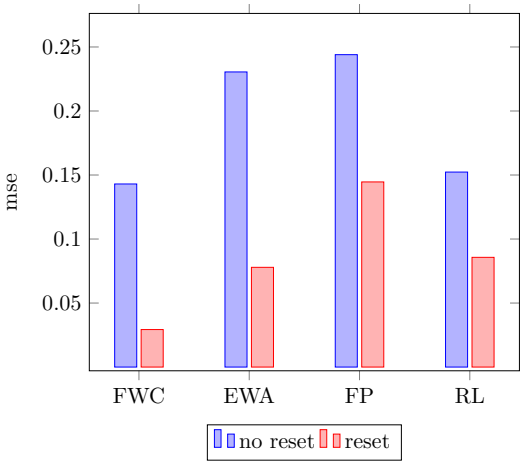


Fig. 11 Comparative evaluation of the mean squared deviation between experimental results and simulated data using Q_2 for the round-by-round average rate of H-play, without and with reset. The panel shows the values for Q_2 computed over 418 datapoints from four treatments

We test for A_5 assuming reset and using $1 - Q_2$; see Appendix A.2 for details. As shown in the fifth row of Fig. 7, the test confirms that only FWC is consistent with A_5.

5 FWC play over other game classes

We distinguish four main classes of similar games: PD (Prisoners' Dilemma), SH (Stag Hunt), CG (Chicken Games), and DE (Prisoners' Delight). The previous section offers a qualitative comparison between FWC and three other learning algorithms over PD and SH. This is based on the extensive empirical evidence collected in Mengel (2018) for PD, and assembled by us for SH.

This section adds three contributions: 1) an exploratory study based on the evidence over CG, where FWC performs better than the competing algorithms; 2) a short argument about the generic equivalence of all algorithms over DE; and 3) some practical suggestions for extending FWC to 3×3 games, tested against a limited dataset. This latter extension is meant as a preliminary exercise: the complexity of categorizing 3×3 games is higher than 2×2 games, and not fully understood yet; see Selten et al. (2003) for an early inquiry into different strategic considerations arising in connection with 3×3 games.

5.1 Chicken games

We sifted the literature and collected five experimental studies that examine a total of six chicken games using the random matching protocol. Detailed information on the six games and relative sources is in Table SM.5. The studies provided us with independent data on the round-by-round average H-rates for three games and on the block-by-block average H-rates for other three games, where a block consists of a consecutive subset of rounds. The actual data are in Table SM.6.

Clearly, the dynamics of the H-rates is more jagged when data are averaged round-by-round than block-by-block. For instance, the left panel of Fig. 12 displays

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

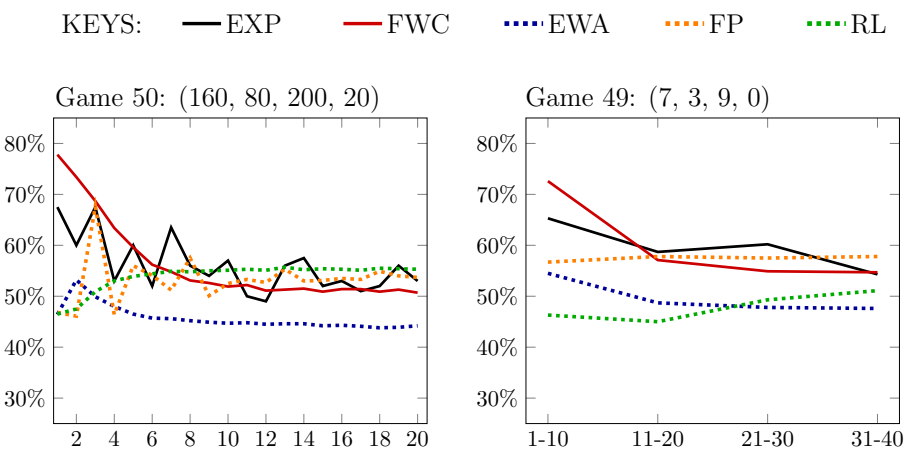


Fig. 12 Round-by-round (left panel) and block-by-block (right panel) average H-rates over two CG games. We depict the experimental results (solid black line) and the simulated *H*-rates averaged over 1000 simulation runs for FWC (red line), EWA (blue line), FP (orange line) and RL (green line)

Table 4 Values of the mean square distances Q_2 for FWC, EWA, FP, RL, CF and NE and the experimental results of the CG and 3×3 games. Each datapoint is based on 1,000 simulation runs

	FWC	EWA	FP	RL	CF	NE
Q_2 (6 CG games)	0.0056	0.0148	0.0087	0.0147	0.0144	0.0069
Q_2 (4 3×3 games)	0.058	0.068	0.05	0.09	0.137	0.122

round-by-round average *H*-rates for a CG, while the right panel displays block-by-block average *H* rates (with block size 10) for another CG. (Analogous displays for all six experiments can be found in Fig. SM.4.) In both games, the equilibrium in mixed strategies prescribes playing *H* with probability 0.6.

The experimental data gravitate close to the equilibrium in mixed strategies. The performance of the algorithms is affected by the initial conditions and by the number of participants. We make no attempt to calibrate parameters, but all the algorithms match the experimental data quite well. FWC and FP appear to track better how the gyrations of the *H*-rates affect the averages. To check this, we define the *block-by-block mean squared error* Q_3 by analogy with the round-by-round mean squared error Q_2; then we compute Q_2 for those treatments where round-by-round average *H*-rates are available and Q_3 for those where only block-by-block average *H*-rates are available, and add them up. Overall, we use 76 experimental datapoints over six different chicken games. A visual summary of the results is in Fig. 13; the actual data are given in Table 4. FWC performs best, followed by FP (Table 5).

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Table 5 Resulting values of the sample tests with softmax versions of FWC, EWA, FP, RL, CF and NE on the 24 PD+SH games.

	FWC	EWA	FP	RL
Q_2 (min)	0.0084	0.0293	0.0296	0.0448
Q_2 (max)	0.0342	0.0819	0.1037	0.1192
Q_2 (median)	0.0258	0.0498	0.0615	0.073
λ^* (min)	4.8	0.02	0.02	0.01
λ^* (max)	6.25	0.1	0.5	1.95
λ^* (median)	5.35	0.05	0.05	0.975

The upper part shows the minimum, maximum and median of Q_2 estimations, the lower part the minimum, maximum and median of corresponding optimal softmax parameter λ^*, each value based on 100 sample experiments per learning rule

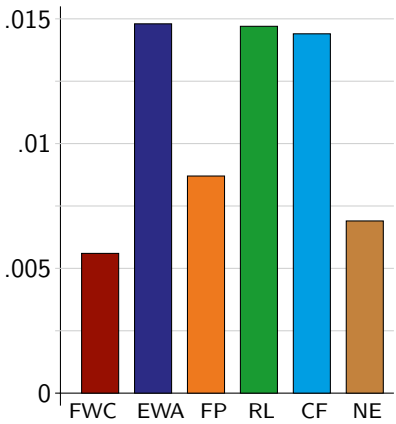


Fig. 13 Comparative evaluation of the mean squared error between experimental results and simulated data (1000 runs per data point) adding Q_2 for the round-by-round and Q_3 for the block-by-block average H-rates, based on 76 data points from six treatments

5.2 Prisoners' delight

Prisoners' Delight collects games where H is the dominant strategy. When both players choose H, they achieve the payoff a, which is either the first-best (if $a > b$) or the second-best (if $a < b$). It easily follows that the only Nash equilibrium is (H, H). Moreover, both FWC and FP predict that only H is played, under any initial condition; instead, EWA and RL predict that the average rate of play for H is increasing over time. All algorithms make similar predictions and, in fact, it is widely believed that most subjects would choose to play H even in a random matching protocol.

5.3 A foray beyond 2×2 games

This paper focuses on symmetric 2×2 games. The graceful suggestion from one referee prompted us to explore how FWC might extend to other games. We focus on symmetric 3×3 games, where $A = [a_{ij}]$ is the payoff matrix for the row player when he plays s_i and the opponent plays s_j. We maintain the assumptions that the payoffs are positive, not all identical, and that the payoffs along the main diagonal are

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

arranged in decreasing order: $a_{11} \geq a_{22} \geq a_{33}$. Actions s_1, s_2, s_3 may be accordingly labeled H (high), M (medium), L (low).

The class of 3×3 games is more complex. There is no comprehensive study that pinpoints a small subset of key descriptors, as Mengel (2018) does for PD games. And it is very likely that different people may pay attention to different features, or that they may use different categorizations. This section argues that a minimalist port of our model for 2×2 games to the class of 3×3 games is possible and that its extension to $n \times n$ games is straightforward. We do not claim that our port is best, or that a successful extension may not require more parameters.

Our plain-vanilla extension of FWC to symmetric 3×3 games is the following. Let $s_i = r(s_j)$ be the best-reply function for the row player; in case of ties, we assume that the best-reply function selects the pure strategy in the "higher" row. Let $M = \max_{i,j} a_{ij}$ be the highest payoff of the game. For each of the three columns of A, we compute the temptation value as:

$$T_i = \frac{\left(\max_{j \neq i} a_{ij}\right) - a_{ii}}{M}$$

The temptation T_i may be positive or negative. When T_i is negative, the best reply is $r(s_i) = s_i$ and greed provides a motivation to stay with s_i. When T_i is positive, the best reply is $r(s_i) \neq s_i$ and greed provides a motivation to move away from s_i towards $r(s_i)$ so as to attain the highest payoff if the opponent plays i. (Note how in 2×2 games a negative temptation about L is equivalent to a positive risk about L.) Along the three temptation values, we compute one efficiency value defined as

$$E = \frac{a_{11} - a_{22}}{M}$$

Efficiency is always positive by construction, because coordination on the best payoff provides a motivation to play s_1.[1]

Any symmetric 3×3 game maps to four extrinsic features (T_1, T_2, T_3, E). The agent's motivations are intrinsic features, whose relative strength is described by four weights (g_1, g_2, g_3, h). The product of the corresponding extrinsic and intrinsic features determines four dispositions $d_{T_i} = g_i \cdot T_i$ (for $i = 1, 2, 3$) and $d_E = h \cdot E$. Then the probability that s_1 (or H) is chosen is

$$P(s_1) = \frac{d_E + (d_{T_1})^- + (d_{T_2})^+ \mathbb{1}(s_1 = r(s_2)) + (d_{T_3})^+ \mathbb{1}(s_1 = r(s_3))}{d_E + |d_{T_1}| + |d_{T_2}| + |d_{T_3}|}$$

where the indicator function $\mathbb{1}(s_i = r(s_s))$ takes value 1 if and only if s_i is the best reply to s_j. Comparing the formula for $P(s_1)$ in a 3×3 game against the formula (1) for $P(H)$ in a 2×2 game, one discerns the common elements: the denominator collects the absolute values for all the dispositions; the numerator uses the positive or

[1] One may consider other efficiency indicators, but we opted for keeping our extension as simple as possible.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

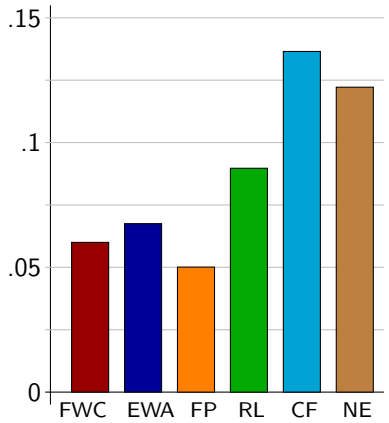


Fig. 14 Comparative evaluation of the round-by-round mean squared error between experimental results and simulated data (1000 runs per data point), based on 88 datapoints from four treatments of 22 rounds each

negative parts of the same dispositions so as to make them active when the disposition favors the strategy. In plain words, the probability of choosing s_1 is increasing in the disposition to efficiency, and in the dispositions to the temptations for which s_1 is the best-reply.

By analogy, ignoring the efficiency motive, the probability that s_2 (or M) is chosen is

$$P(s_2) = \frac{(d_{T_1})^+ \mathbb{1}(s_2 = r(s_1)) + (d_{T_2})^- + (d_{T_3})^+ \mathbb{1}(s_2 = r(s_3))}{d_E + |d_{T_1}| + |d_{T_2}| + |d_{T_3}|}$$

and $P(s_3)$ is analogously defined.

The update rule is based on the last observed opponent's action, similarly as above. At time t, the current strengths $(g_1, g_2, g_3, h)_i$ for a player i are updated using the rule

$$(g_1', g_2', g_3', h')_i = \begin{cases} (g_1 + \gamma_i,\ g_2,\ g_3,\ h)_i & \text{if the last opponent's choice is } (s_1) \\ (g_1,\ g_2 + \gamma_i,\ g_3,\ h)_i & \text{if the last opponent's choice is } (s_2) \\ (g_1,\ g_2,\ g_3 + \gamma_i,\ h)_i & \text{if the last opponent's choice is } (s_3) \\ (g_1,\ g_2,\ g_3,\ h + \gamma_i)_i & \text{if the last opponent's choice is } (s_1) \text{ and } (s_1, s_1) \text{ gives the highest possible payoff} \end{cases}$$

and afterwards they are divided by $(1 + \gamma_i)$ to renormalize their sum to 1. The learning rate for agent i is set as $\gamma_i = 1 - \max_k P_i(s_k)$.

This plain-vanilla extension of FWC is meant only as a proof-of-concept. We could not amass enough experimental evidence to subject it to a full comparative analysis, and we do not claim that it would pass muster. As a first exploratory step, we looked at the data from Cooper et al. (1990) about 4 identical 3×3 symmetric coordination games, based on a random matching protocol with an

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

additional restriction. See Table SM.7 for detailed information on the games and Table SM.8 for the round-by-round data. Figure 14 summarizes the mean squared error Q_2 for the round-by-round average H-, M-, and L-rates against simulated data. Encouragingly, FWC's performance is second to FP's.

6 Concluding remarks

The main goal of this paper is to advance a category-based approach to the study of learning across *similar* games played against different opponents. We implemented plain-vanilla versions for ours and three major competing algorithms, comparing their performances against experimental data from others' studies.

The available evidence supports the efficacy of a category-based approach to describe and predict aggregate play. On the other hand, because we did not try to calibrate the specific models to the empirical evidence, there remains a lingering question about the robustness of our comparisons.[2]

Section A.3 in the appendix offers three independent checks that corroborate the soundness of our results. Specifically: (a) we replace the self-adjusting learning rate $\gamma = 1 - \max(P_i(H), P_i(L))$ in FWC with a fixed learning rate γ in (0, 1); (b) we make the update rule for FWC depend on a longer memory than the last period; (c) we implement a one-parameter softmax version for each of the four learning algorithms and, after calibrating the parameter on half of the experimental evidence, we test the optimized model on the second half.

To conclude, we suggest a novel approach where agents learn to categorize games and tend to play the same action for games placed in the same category. We demonstrate its potential by implementing a robust (parameter-free) model. When tested over a large body of independent evidence for identical games, our model fits the empirical data better than the current major competitors. Similarly, it provides a superior match for the independent evidence over similar games.

We emphasize that the major concern of the paper is to argue that a sound modeling framework for learning over similar games is within reach. Future experimental studies may expand on this perspective and provide more data about how agents' behave when facing a stream of similar games.

Online Appendix The online appendix, including Supplementary Material, is available at https://doi.org/10.1007/s10683-021-09742-7.

Acknowledgements We thank audiences at LEG 2019 and SAET 2019 for their comments. This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 732942. Furthermore, RM has received funding from the Polish National Agency for Academic Exchange NAWA (Narodowa Agencja Wymiany Akademickiej), Grant No. PPN/ULM/2019/1/00222.

[2] We thank an attentive referee for pointing this out.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Author contributions MLC and RM planned the study and built the computational model. RM ran the simulations. MLC and RM collected the experimental data. MLC and RM analysed the data and wrote the manuscript.

Data availability All data derived from others' studies are listed as Supplementary Material in Tables SM.1, SM.2, SM.3 and SM.4.

Code availability The algorithms are implemented and tested using *Python*. The project is accessible at: https://figshare.com/projects/Feature_weighted_categorized_play_across_symmetric_games/127253.Code doi: https://doi.org/10.6084/m9.figshare.17081408.

Declarations

Conflict interest The authors declare no competing interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Ahn, T., Ostrom, E., Schmidt, D., Shupp, R., & Walker, J. (2001). Cooperation in PD games: Fear, greed, and history of play. *Public Choice, 106,* 137–155.

Andreoni, J., & Miller, J. H. (1993). Rational cooperation in the finitely repeated prisoners dilemma: Experimental evidence. *Economic Journal, 103*(418), 570–585.

Boylan, R. T., & El-Gamal, M. (1993). Fictitious play: A statistical study of multiple economic experiments. *Games and Economic Behavior, 5*(2), 205–222.

Camerer, C. F., & Ho, T.-H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica, 67*(4), 827–874.

Chmura, T., Goerg, S. J., & Selten, R. (2012). Learning in experimental 2×2 games. *Games and Economic Behavior, 76*, 44–73.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1990). Selection criteria in coordination games: Some experimental results. *The American Economic Review, 80*(1), 218–233.

Crawford, V. P. (2002). Introduction to experimental game theory. *Journal of Economic Theory, 104*, 1–15.

Daskalova, V., & Vriend, N. J. (2021). Learning frames. *Journal of Economic Behavior & Organization*, *191*, 78–96.

Davis, L. H. (1977). Prisoners, paradox, and rationality. *American Philosophical Quarterly, 14*, 319–327.

Devetag, G., & Warglien, M. (2008). Playing the wrong game: An experimental analysis of relational complexity and strategic misrepresentation. *Games and Economic Behavior, 62*, 364–382.

Duffy, J., & Fehr, D. (2018). Equilibrium selection in similar repeated games: Experimental evidence on the role of precedents. *Experimental Economics, 21*, 573–600.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88*(4), 848–881.

Feltovich, N. (2000). Reinforcement-based vs. belief-based learning models in experimental asymmetric-information games. *Econometrica, 68*(3), 605–641.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Flache, A., & Macy, M. W. (2006). Learning and framing in social exchange. In B. B. Detlef Fetchenhauer, Andreas Flache, & S. Lindenberg (Eds.), *Solidarity and Prosocial Behavior.* New York: Springer.

Gibbons, R., LiCalzi, M., & Warglien, M. (2021). What situation is this? Coarse cognition and behavior over a space of games. *Strategy Science, 6*, 124–140.

Gilboa, I., & Schmeidler, D. (1995). Case-based decision theory. *Quarterly Journal of Economics, 110*, 605–639.

Gilboa, I., & Schmeidler, D. (2001). *A theory of case-based decisions*. Cambridge (UK): Cambridge University Press.

Grimm, V., & Mengel, F. (2012). An experiment on learning in a multiple games environment. *Journal of Economic Theory, 147*, 2220–2259.

Hagenbach, J., & Koessler, F. (2020). Cheap talk with coarse understanding. *Games and Economic Behavior, 124*, 105–121.

Haruvy, E., & Stahl, D. O. (2012). Between-game rule learning in dissimilar symmetric normal-form games. *Games and Economic Behavior, 74*(1), 208–221.

Heller, Y., & Winter, E. (2016). Rule rationality. *International Economic Review, 57*(3), 997–1026.

Ho, T.-H., Camerer, C. F., & Chong, J.-K. (2007). Self-tuning experience weighted attraction learning in games. *Journal of Economic Theory, 133*, 177–198.

Hofstadter, D. R. (1983). Metamagical themas. *Scientific American, 248*, 14–29.

Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory, 123*, 81–104.

Jehiel, P. (2020). Analogy-based expectation equilibrium and related concepts: Theory, applications, and beyond. 6th October 2020.

Knez, M., & Camerer, C. F. (2000). Increasing cooperation in prisoners dilemmas by establishing a precedent of efficiency in coordination games. *Organizational Behavior and Human Decision Processes, 82*(2), 194–216.

Lensberg, T., & Schenk-Hoppé, K. R. (2021). Cold play: Learning across bimatrix games. *Journal of Economic Behavior and Organization, 185*, 419–441.

LiCalzi, M., & Mühlenbernd, R. (2019). Categorization and cooperation across games. *Games*, *10*.

Mengel, F. (2012). Learning across games. *Games and Economic Behavior, 74*(2), 601–619.

Mengel, F. (2018). Risk and temptation: A meta-study on prisoners dilemma games. *Economic Journal, 128*(616), 3182–3209.

Mullainathan, S., Schwartzstein, J., & Shleifer, A. (2008). Coarse thinking and persuasion. *Quarterly Journal of Economics, 123*, 577–619.

Peysakhovich, A., & Rand, D. G. (2016). Habits of virtue: Creating norms of cooperation and defection in the laboratory. *Management Science, 62*(3), 631–647.

Rankin, F. W., Van Huyck, J. B., & Battalio, R. C. (2000). Strategic similarity and emergent conventions: Evidence from similar stag hunt games. *Games and Economic Behavior, 32*, 315–337.

Rapoport, A. (1966). *Two-Person Game Theory*. Ann Arbor: University of Michigan Press.

Rapoport, A., & Guyer, M. (1966). A taxonomy of 2×2 games. *General Systems, 11*, 203–214.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior, 8*, 164–212.

Rusch, H., & Luetge, C. (2016). Spillovers from coordination to cooperation: Evidence for the interdependence hypothesis? *Evolutionary Behavioral Science, 10*(4), 284–296.

Samuelson, L. (2001). Analogies, adaptation, and anomalies. *Journal of Economic Theory, 97*, 320–366.

Schmidt, D., Shupp, R., Walker, J., Ahn, T., & Ostrom, E. (2001). Dilemma games: Game parameters and matching protocols. *Journal of Economic Behavior and Organization, 46*, 357–377.

Selten, R. (1998). Axiomatic characterization of the quadratic scoring rule. *Experimental Economics, 1*, 43–62.

Selten, R., Abbink, K., Buchta, J., & Sadrieh, A. (2003). How to play 3×3 games. A strategy method experiment. *Games and Economic Behavior, 45*, 19–37.

Steiner, J., & Stewart, C. (2008). Contagion through learning. *Theoretical Economics, 3*, 431–458.

Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: The MIT Press.

Van Huyck, J. B., & Stahl, D. O. (2018). Conditional behavior and learning in similar stag hunt games. *Experimental Economics, 21*, 513–526.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

https://doi.org/10.1007/s10683-021-09742-7 Published online by Cambridge University Press