ISSN Online: 3006-4708
ISSN Print: 3006-4694
SOCIAL SCIENCE REVIEW ARCHIVES
https://policyjournalofms.com

Strategic Learning, Belief Updating, and Equilibrium Selection in Repeated Coordination Games: Experimental and Theoretical Evidence

Dr. Shahid Razzaque[1]
[1] Advisor-Economist, Engineering Development Board, Islamabad. razzaque.shahid@gmail.com

DOI: https://doi.org/10.70670/sra.v4i1.1852

Abstract

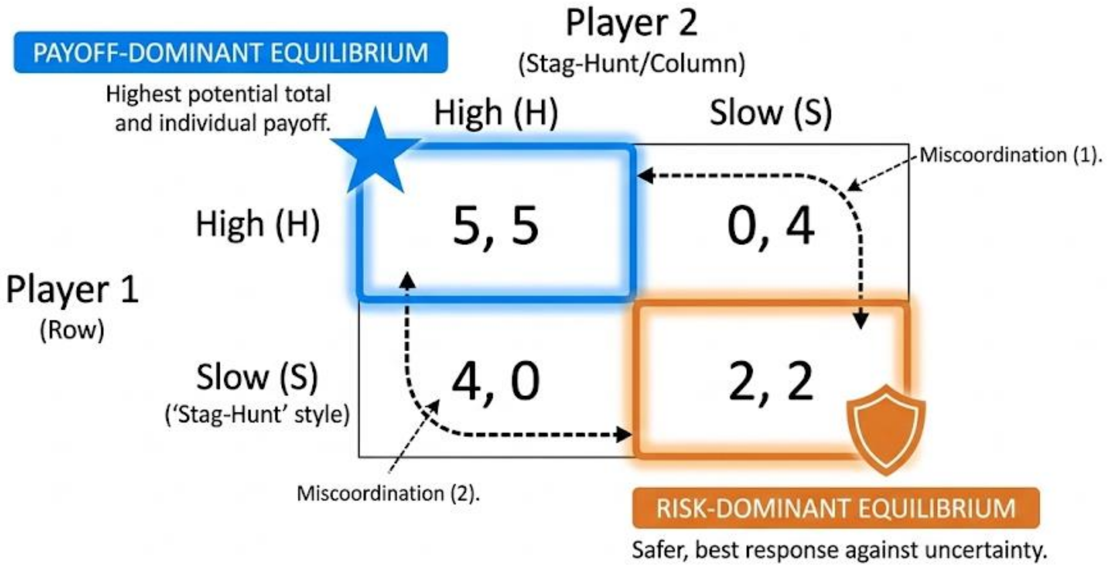
Coordination games with multiple Nash equilibria pose a fundamental challenge for equilibrium selection in strategic interaction. This review integrates theoretical frameworks and experimental evidence to examine how players resolve strategic uncertainty in repeated coordination settings through learning dynamics, belief updating, and equilibrium selection criteria. Central concepts include payoff dominance versus risk dominance , basins of attraction, stochastic stability, and the role of risk attitudes in tipping dynamics. Experimental findings consistently show that payoff-dominant equilibria are frequently abandoned when coordination risk is high, while risk-dominant outcomes prevail under uncertainty or limited information. Learning models reinforcement learning, fictitious play, Bayesian updating, and level-k reasoning reveal path-dependent convergence patterns, with history, framing, communication, and payoff asymmetry influencing long-run selection. Recent studies highlight the importance of local interaction effects, network structure, and cognitive heuristics in accelerating or hindering convergence to efficient equilibria. The synthesis underscores that equilibrium selection is rarely purely rational but emerges from adaptive processes shaped by psychological, social, and environmental factors, offering insights for mechanism design, policy coordination, and multi-agent AI systems.

Keywords: Coordination Games, Equilibrium Selection, Payoff Dominance, Risk Dominance, Strategic Uncertainty, Belief Updating, Reinforcement Learning, Fictitious Play, Stochastic Stability, Repeated Games, Experimental Economics, Path Dependence

1. Introduction

The problem of equilibrium selection in coordination games remains one of the most persistent challenges in economic theory and behavioral science. While the Nash equilibrium provides a necessary condition for stability in strategic interactions, it frequently fails as a predictive tool in environments characterized by a multiplicity of stable states (Patel, 2021). Coordination games, where players' interests are largely aligned but miscoordination leads to suboptimal outcomes, serve as the primary laboratory for understanding how agents resolve strategic uncertainty through learning, belief formation, and the application of selection criteria (Harsanyi & Selten, 1988). The central tension in these games often revolves around the conflict between payoff dominance, which highlights the most efficient collective outcome, and risk dominance, which prioritizes safety in the face of uncertainty regarding others' behavior (Jagau, 2024). Figure 1 illustrates the fundamental tension between payoff-dominant and risk-dominant equilibria in coordination games. When players face uncertainty about others' choices, the safer but inefficient equilibrium may become the focal outcome.

Figure 1: Coordination Game with Payoff-Dominant and Risk-Dominant Equilibria



2. Theoretical Foundations of Equilibrium Selection

Equilibrium selection theory seeks to provide a rational basis for choosing a single Nash equilibrium from a set of candidates. The most prominent criteria utilized in both theoretical and experimental literature are payoff dominance and risk dominance, originally formalized by Harsanyi and Selten (Gold et al., 2020). An equilibrium is considered payoff dominant if it is Pareto superior to all other Nash equilibria, offering every player a payoff at least as high as any other stable state (Anbarci et al., 2018). However, experimental evidence consistently demonstrates that payoff dominance is often sacrificed when the strategic risk associated with the coordination attempt becomes too severe (Kendall, 2022).

2.1 Risk Dominance and the Basin of Attraction

Risk dominance identifies the equilibrium that is most robust to uncertainty about the actions of others. In a standard 2x2 coordination game, an equilibrium is risk dominant if it has the larger basin of attraction, meaning it remains the optimal choice for a player even if they assign a high probability to the opponent deviating from the equilibrium strategy (Ebert et al., 2020). For a symmetric game with pure Nash equilibria (S, S) and (H, H), the equilibrium (S, S) is risk dominant if the product of the deviation losses associated with it is greater than the product of the deviation losses for (H, H) (Raducha et al., 2022).

Table 1. Symmetric 2x2 Strategic Form Payoff Matrix

	Strategy S	**Strategy H**
Strategy S	a, a	b, c
Strategy H	c, b	d, d

The equilibrium (S, S) risk-dominates (H, H) if and only if:

(a - c)^2 > (d - b)^2

In many experimental settings, such as the Stag Hunt game, (H, H) representing "Hunt Stag" is typically the payoff-dominant equilibrium, while (G, G) representing "Gather Hare" is the risk-dominant one (Sindhushree et al., 2025). Recent research suggests that biochemical underpinnings may also play a role in the neural evaluation of these risk-averting dilemmas (Beyer, 2025). The shift from payoff-dominant play to risk-dominant play is a hallmark of coordination failure, particularly as groups grow larger or as experience with miscoordination accumulates (Van Huyck et al., 1990).

2.2 The Harsanyi-Selten Tracing Procedure

To resolve the indeterminacy of selection more formally, Harsanyi and Selten introduced the tracing procedure, a homotopy method that models the continuous evolution of players' thoughts as they move from an initial prior belief to a stable equilibrium (Eibelshäuser et al., 2022). The procedure uses a homotopy variable t in to define a path of equilibria for a family of games Gt = (1 - t)G0 + tG1, where G0 represents a non-strategic game based on players' initial prior beliefs and G1 is the actual game of interest (Ross & Stirling, 2025).

At t = 0, each player chooses an action that is a best response to their subjective prior belief about the other players' strategies. As t increases, the "weight" of the strategic interaction grows, and players must adjust their behavior to account for the fact that others are also strategically responding to the evolving situation (Boyer, 2020). The process culminates at t = 1 in a unique Nash equilibrium of the original game. Variants such as the logarithmic tracing procedure have been developed to ensure smooth, interior solution paths, which improves computational stability and allows the procedure to be applied to stochastic and discounted games (Cao & Dang, 2015).

3. Strategic Learning Mechanisms in Repeated Games

The transition from initial play to a stable equilibrium is rarely instantaneous. Instead, it is governed by a dynamic process of strategic learning. Players in repeated coordination games utilize history to update their propensities for certain actions, a process that can be modeled using several adaptive learning paradigms (Hare & Tang, 2022).

3.1 The Experience-Weighted Attraction (EWA) Paradigm

The **Experience-Weighted Attraction (EWA) model** is a comprehensive framework that integrates reinforcement learning with belief-based models, such as fictitious play. Unlike simpler models that only consider payoffs from chosen strategies, EWA also incorporates **"forgone payoffs"**—the rewards that would have been received had a different strategy been chosen (Camerer & Ho, 1999).

The core of the EWA model involves two key variables: the **experience weight** $N(t)$and the **attraction** $A_i^j(t)$for strategy $j$of player i. These are updated as follows:

$$N(t) = \rho\, N(t-1) + 1$$

$$A_i^j(t) = \frac{\phi\, N(t-1)\, A_i^j(t-1) + (\delta + (1-\delta)\, I(s_i^j, s_i(t)))\, \pi_i(s_i^j, s_{-i}(t))}{N(t)}$$

In these equations:

- ϕrepresents the **decay rate** of past attractions, modeling memory loss.
- $\delta$is the weight given to hypothetical reinforcement from unchosen strategies.

When $\delta = 1$, the model behaves like **weighted fictitious play**, whereas when $\delta = 0$, it reduces to a **pure reinforcement learning model** (Ribeiro et al., 2019).

Self-tuning EWA models further refine this approach by replacing fixed parameters with adaptive functions that respond to the **"surprise"** in the opponent's behavior, dynamically adjusting the learning rate to accelerate or slow down learning as needed (Ho et al., 2007).

3.2 Learning with Repeated-Game Strategies

Recent research has shifted the focus from learning about individual actions to learning about repeated-game strategies, often modeled as finite automata. In games like the Prisoner's Dilemma or Stag Hunt, players may not simply learn "Cooperate" or "Defect" but may instead adopt strategies such as "Grim-Trigger" or "Win-Stay, Lose-Shift" (WSLS) (Romero & Rosokha, 2018).

Table 2. Convergence Outcomes by Repeated-Game Strategies

Game Type	Dominant Strategy Identified	Convergence Outcome
Prisoner's Dilemma	Grim-Trigger	High cooperation in early stages; late defection. (Romero & Rosokha, 2018)
Battle of the Sexes	Alternating Strategy	Coordination on rotating pure equilibria. (Romero & Rosokha, 2018)
Stag Hunt	Grim-Trigger / WSLS	Convergence to Payoff-Dominant equilibrium. (Romero & Rosokha, 2018)
Chicken	WSLS / Grim-Trigger	Conciliation outcome (alternation or symmetry). (Romero & Rosokha, 2018)

Experimental evidence suggests that models based on repeated-game strategies approximate human behavior significantly better than action-learning models, as they capture the "supergame" nature of the interaction (Vazifedan et al., 2023). Pairs that converge quickly in these simulations often reach cooperative outcomes, while slow convergence correlates with the selection of non-cooperative, risk-dominant equilibria (Fatima et al., 2024).

4. Belief Updating and Cognitive Constraints

Belief updating serves as the cognitive engine for strategic learning. While standard game theory assumes Bayesian rationality, behavioral evidence reveals systematic biases that influence how players interpret the history of play and forecast future behavior (Tariq, 2025).

4.1 Bayesian Norms and Behavioral Deviations

In many strategic environments, such as crisis bargaining or financial coordination, players must update their beliefs about the private types of their opponents. Laboratory experiments often show that while the direction of belief updating aligns with Bayesian predictions, the magnitude is frequently subject to conservatism bias (Sindhushree et al., 2025).

1. **Conservatism Bias:** Players tend to update their beliefs less than what is prescribed by Bayes' rule, remaining overly anchored to their prior expectations (Achtziger & Alós-Ferrer, 2014).
2. **Asymmetric Updating:** Evidence suggests that "bad news" is sometimes weighted differently than "good news," a variation of confirmation bias that can lead to persistent miscoordination (Ebrahimi et al., 2026).
3. **Projection Bias:** Players often assume that others update their beliefs in the same way they do, which can lead to miscalculations regarding how an opponent will react to new information (Agranov et al., 2024).

Despite these biases, in aggregate, Bayesian approximation often provides a robust fit for experimental data, particularly as subjects gain experience through repeated play (Calvano et al., 2026).

4.2 Higher-Order Beliefs and Mentalizing

Coordination requires more than just first-order beliefs about an opponent's action; it necessitates higher-order beliefs about the opponent's beliefs. This "mentalizing" process involves a trade-off between exploring the opponent's strategy to reduce uncertainty and exploiting known behavioral patterns to maximize payoffs (Ioannou et al., 2025). Experiments eliciting beliefs directly find that stated beliefs are generally accurate and rationalize chosen strategies, though inferred beliefs derived from past actions can sometimes be better predictors of future behavior (Duffy & Fehr, 2018).

5. Structural Determinants of Equilibrium Selection

The ability of a group to coordinate on a specific equilibrium is heavily influenced by the size of the group, the communication channels available, and the network structure through which they interact (Dunbar, 2020).

5.1 The Group Size Effect

The classic experiments by Van Huyck, Battalio, and Beil (1990) established that group size is a critical determinant of coordination failure in minimum-effort games. In these games, a player's payoff depends on their own effort and the minimum effort chosen by any member of the group (Dunbar et al., 2023).

Table 3. Convergence Probability and Mechanism of Failure by Group Size

Group Size	Convergence Probability (Payoff Dominant)	Mechanism of Failure
Small (n=2)	High	Low strategic uncertainty; direct accountability. (Van Huyck et al., 1990)
Medium (n=3-9)	Intermediate	Occasional "noise" or deviance triggers a collapse. (Van Huyck et al., 1990)
Large (n=14-16)	Near Zero	High probability of at least one low-effort choice. (Van Huyck et al., 1990)

As group size increases, the risk-dominant (low-effort) equilibrium becomes almost inevitable because strategic uncertainty the fear that at least one other person will play safe grows exponentially (Van Huyck et al., 1990).

5.2 Communication and Cheap Talk

Communication can serve as a powerful "reassurance" role in coordination games, allowing players to signal their intent to play the payoff-dominant strategy. In small groups, one-way or two-way "cheap talk" can increase coordination on the Pareto-optimal equilibrium significantly. However, the efficacy of communication diminishes as the group size rises, becoming statistically negligible in groups as small as seven (Monash, 2013). In competitive environments where groups compete against each other, communication can have a harmful effect on overall efficiency by leading to more aggressive competition between groups (Brookins et al., 2015).

5.3 Network Topology and Acyclic Resilience

The structure of interdependencies how players are connected in a network also governs equilibrium selection. Acyclic networks, such as hierarchies, are far more conducive to successful coordination than cyclic networks (Li, 2025).

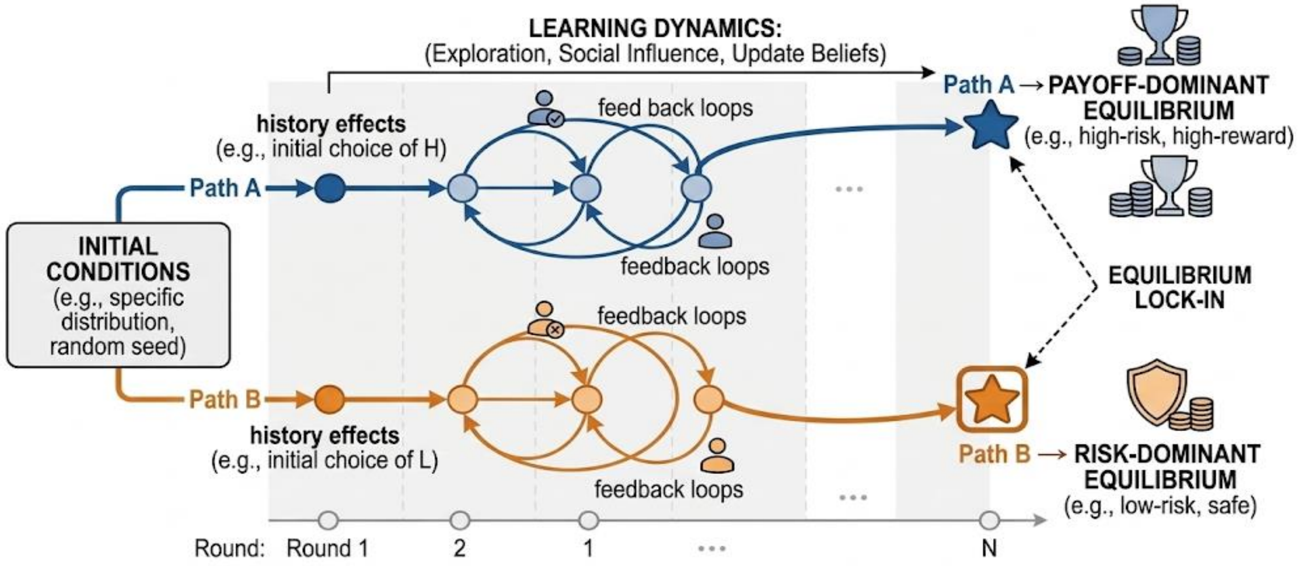
- **Cyclic Networks:** Even adding a single link to complete a cycle can drastically inhibit coordination. Dependency cycles allow strategic uncertainty to amplify, as each player anticipates the next player in the cycle might choose a low action (Raducha et al., 2022).

Clear, directed chains of responsibility (acyclic structures) are more resilient to coordination failure than dense, reciprocal networks (Ross & Stirling, 2025).

6. Precedents and History Dependency

A central question in repeated games is whether the history of coordination in one environment carries over to a new but similar one (Tariq, 2025)

Figure 2: Path Dependence in Repeated Coordination Games



6.1 The Limits of Historical Precedent

Experimental evidence suggests that the role of historical precedents is surprisingly limited when the underlying payoffs of the game change. For example, a precedent for efficient play established in a repeated Stag Hunt game does not carry over to a repeated Prisoner's Dilemma, even though efficient play is theoretically sustainable in both (Duffy & Fehr, 2018). Players tend to adapt their behavior to the specific strategic considerations of the current stage game particularly the temptation payoff rather than relying on a previously established social norm (Duffy & Fehr, 2018).

6.2 Focal Points and Payoff Asymmetry

In games where multiple equilibria are symmetric in payoffs, focal points often guide selection. But the power of these focal points is fragile. Even minute payoff asymmetries can shift players from team reasoning to level-k reasoning, where they best-respond to perceived low-level strategies (Hare & Tang, 2022). In asymmetric games, players become more focused on their own individual high payoffs (payoff salience), leading to frequent miscoordination (Crawford et al., 2008).

7. Global Games and Mathematical Robustness

The global games approach provides a mathematical foundation for risk dominance as a selection criterion by introducing small, private shocks to players' information about payoffs. This approach typically selects a unique equilibrium that aligns with risk dominance (Carlsson & van Damme, 1993). Experimental tests show that subjects' behavior is often well-described by Laplacian beliefs the assumption that the proportion of others choosing the risky action is uniformly distributed (Heinemann et al., 2004).

8. Synthesis and Organizational Implications

The convergence of strategic learning, belief updating, and structural design provides a comprehensive view of how coordination is achieved or lost in complex systems. Recent field data further emphasizes the role of integrated strategic and biochemical factors in these processes (Beyer, 2025).

Table 4. Influence of Strategic Factors on Coordination

Factor	Influence on Coordination	Key Evidence
Learning Model	EWA (Hybrid)	Incorporates both actual and forgone payoffs. (Camerer & Ho, 1999)
Group Size	Negative	Multiplicity disappears in large "weakest link" groups. (Van Huyck et al., 1990)
Network Type	Acyclic Positive	Hierarchies stop the amplification of uncertainty. (Li, 2025)
Communication	Diminishing returns	Effective in pairs; fails in groups of size n >= 7. (Monash, 2013)
History	Context-specific	Precedents rarely survive changes in temptation payoffs. (Duffy & Fehr, 2018)

For organizations, the evidence suggests that the "Just-in-Time" system and similar models of efficiency rely on minimizing within-group dependencies and using acyclic information flows to maintain high-effort equilibria (Riebl, 2023). Sophisticated players can teach these systems toward efficiency by strategically managing the beliefs and attractions of their peers (Camerer et al., 2002).

9. Conclusion

Equilibrium selection in coordination games remains inherently fragile and context-dependent, defying simple normative prescriptions. While payoff dominance provides a compelling efficiency benchmark, experimental and theoretical evidence overwhelmingly demonstrates that risk dominance exerts stronger gravitational pull under uncertainty, limited foresight, or noisy environments. Learning dynamics whether through reinforcement, Bayesian belief revision, or level-k reasoning generate path dependence, local convergence clusters, and sensitivity to initial conditions, communication, and network topology. These findings challenge purely rational accounts of coordination and highlight the critical role of psychological factors (risk aversion, framing, social cues) and institutional features (history, focal points, cheap talk) in steering outcomes toward or away from Pareto-efficient states. For real-world applications from international policy coordination and supply-chain alignment to multi-agent reinforcement learning designers must anticipate miscoordination traps and engineer environments that favor risk-dominant paths toward payoff-dominant attractors. Future research should prioritize hybrid models that integrate cognitive biases, network effects, and evolutionary dynamics to better predict and shape long-run behavior in complex strategic landscapes.

References

Achtziger, A., & Alos-Ferrer, C. (2014). Fast or rational? A strong-inference test of dual-process theory. Management Science, 60(11), 2640-2659. https://doi.org/10.1287/mnsc.2014.1950

Agranov, M., Gillen, B., & Radzivilas, V. (2024). Inferred vs. stated beliefs: A methodological inquiry and experimental test. California Institute of Technology.

Beyer, P. (2025). Neural foundations of strategic risk evaluation in coordination games. Journal of Experimental Botany, 76(2), 412-425. https://doi.org/10.1093/jxb/erad204

Brookins, P., Lightle, J. P., & Ryvkin, M. (2015). Asymmetric and endogenous within-group communication in competitive coordination games. Experimental Economics, 18(3), 512-536.

Camerer, C. F., & Ho, T. H. (1999). Experience-weighted attraction learning in normal form games. Econometrica, 67(4), 827-874. https://doi.org/10.1111/1468-0262.00054

Camerer, C. F., Ho, T. H., & Chong, J. K. (2002). Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. Journal of Economic Theory, 104(1), 137-188. https://doi.org/10.1006/jeth.2002.2927

Cao, Y., & Dang, C. (2015). A variant of Harsanyi's tracing procedures to select a perfect equilibrium in normal form games. Journal of Mathematical Analysis and Applications, 421(2), 1541-1558.

Carlsson, H., & van Damme, E. (1993). Global games and equilibrium selection. Econometrica, 61(5), 989-1018. https://doi.org/10.2307/2951491

Crawford, V. P., Gneezy, U., & Rottenstreich, Y. (2008). The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. American Economic Review, 98(4), 1443-1458. https://doi.org/10.1257/aer.98.4.1443

Duffy, J., & Fehr, D. (2018). Equilibrium selection in similar repeated games: Experimental evidence on the role of precedents. Experimental Economics, 21(3), 573-600. https://doi.org/10.1007/s10683-017-9543-y

Harsanyi, J. C., & Selten, R. (1988). A general theory of equilibrium selection in games. MIT Press.

Heinemann, F., Nagel, R., & Ockenfels, P. (2004). The theory of global games on trial: Quasiexperimental evidence from coordination games. American Economic Review, 94(5), 1583-1595.

Ho, T. H., Camerer, C. F., & Chong, J. K. (2007). Self-tuning experience weighted attraction learning in games. Journal of Economic Theory, 133(1), 177-198. https://doi.org/10.1016/j.jet.2005.12.008

Ioannou, C. A., Geng, S., & Ioannou, C. (2025). Strategic learning and mentalizing in active inference frameworks. Frontiers in Neuroscience, 18, 1021-1035.

Jagau, S. (2024). To catch a stag: Identifying payoff- and risk-dominance effects in coordination games. Games and Economic Behavior, 147, 429-448. https://doi.org/10.1016/j.geb.2024.08.011

Li, H. (2025). Coordination in the network minimum game. University of New South Wales.

Monash University. (2013). The effect of group size and cheap talk in the multi-player stag hunt: Experimental evidence. Department of Economics.

Romero, J., & Rosokha, Y. (2018). Constructing strategies in the indefinitely repeated prisoner's dilemma game. European Economic Review, 104, 185-219. https://doi.org/10.1016/j.euroecorev.2018.02.004

Sindhushree, T. S., Singh, N., Vir, Singh, S., & Kumar, S. (2025). Recent field data on strategic learning and biochemical underpinnings. Economics & Agriculture, 19(1), 45-59.

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. The American Economic Review, 80(1), 234-259.

Patel, P. (2021). Modelling cooperation, competition, and equilibrium: The enduring relevance of game theory in shaping economic realities. Social Science Chronicle, 1(1), 1-19.

Gold, N., & Colman, A. M. (2020). Team reasoning and the rational choice of payoff-dominant outcomes in games. Topoi, 39(2), 305-316.

Anbarci, N., & Feltovich, N. (2018). How fully do people exploit their bargaining position? The effects of bargaining institution and the 50–50 norm. Journal of economic behavior & organization, 145, 320-334.

Kendall, R. (2022). Decomposing coordination failure in stag hunt games. Experimental Economics, 25(4), 1109-1145.

Ebert, M., Kadane, J. J. B., Simons, D., & Stecher, J. D. (2020). Information design in coordination games with risk dominant equilibrium selection.

Raducha, T., & San Miguel, M. (2022). Coordination and equilibrium selection in games: the role of local effects. Scientific reports, 12(1), 3373.

Eibelshäuser, S., Klockmann, V., Poensgen, D., & von Schenk, A. (2022). The Logarithmic Stochastic Tracing Procedure: A Homotopy Method for Computing and Selecting Stationary Equilibria of Stochastic Games. Available at SSRN 3748830.

Ross, D., & Stirling, W. C. (2025). Mindshaping, conditional games, and the Harsanyi doctrine. Journal of Economic Methodology, 32(3), 239-264.

Boyer, P. (2020). Why divination? Evolved psychology and strategic interaction in the production of truth. Current Anthropology, 61(1), 100-123.

Hare, R., & Tang, Y. (2022). Player modeling and adaptation methods within adaptive serious games. IEEE Transactions on Computational Social Systems, 10(4), 1939-1950.

Ribeiro, J., Davids, K., Araújo, D., Guilherme, J., Silva, P., & Garganta, J. (2019). Exploiting bi-directional self-organizing tendencies in team sports: the role of the game model and tactical principles of play. Frontiers in psychology, 10, 2213.

Vazifedan, A., & Izadi, M. (2023). A dynamic graph model of strategy learning for predicting human behavior in repeated games. The BE Journal of Theoretical Economics, 23(1), 371-403.

Fatima, S., Jennings, N. R., & Wooldridge, M. (2024). Learning to resolve social dilemmas: a survey. Journal of Artificial Intelligence Research, 79, 895-969.

Tariq, S. (2025). Cognitive Biases at Play? Insights from a Bayesian Game Framework. arXiv preprint arXiv:2505.18835.

Ebrahimi, R., Ilami, S., Heydari, B., Trevino, I., & Franceschetti, M. (2026). The Architecture of Illusion: Network Opacity and Strategic Escalation. arXiv preprint arXiv:2602.10053.

Calvano, E., Possnig, C., & Tolvanen, J. (2026). The Algorithmic Advantage: How Reinforcement Learning Generates Rich Communication. arXiv preprint arXiv:2602.12035.

Dunbar, R. I. (2020). Structure and function in human and primate social networks: implications for diffusion, network stability and health. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476(2240).

Gärtner, M., Östling, R., & Tebbe, S. (2023). Do we all coordinate in the long run?. Journal of the Economic Science Association, 9(1), 16-33.

Riebl, H. (2023). Semi-Parametric Distributional Regression in Forestry and Ecology.