Introspection dynamics in asymmetric multiplayer games
Reply to reviewers

Marta C. Couto[1*] and Saptarshi Pal[1*]

[1]Max Planck Research Group Dynamics of Social Behavior, Max Planck Institute for Evolutionary Biology, 24306 Ploen, Germany

[*] co-first authors

Summary

We thank the reviewers for their summaries and comments. We believe that the questions raised are pertinent and they led to the manuscript's improvement. Below we reply to the reviewers' specific comments. We also include a new version of the manuscript where the changes are highlighted. Mainly, we add three new propositions and discuss the connection of introspection dynamics to the logit-response model. Additionally, we reviewed the whole text and added a few new references.

Replies

Reviewer 1

Summary

In this paper, the authors explore a model for social learning in an n-player, asymmetric game called introspection dynamics, in which players play a repeated game and a single player is chosen at random to have the opportunity to revise their action for the next round. The agent revising their strategy picks a new possible action at random, and then chooses to adopt that action with a probability depending both on the difference in payoffs between the player's current strategy and the new possible strategy (given the current actions of the other players in the population) and on a parameter describing the strength of selection depending on payoff differences. The authors formulate a Markov chain describing how the actions of the n players evolve under introspection dynamics, showing that the Markov chain is irreducible and characterizing the stationary distribution achieved as the long-time behavior for introspection dynamics. The authors also study two special cases of asymmetric additive games and n-player, two-strategy symmetric games, calculating explicit stationary distribution in these cases and illustrating the strategic composition at the stationary distribution for example games of a linear public goods game with heterogeneous costs and benefits and for a symmetric nonlinear public goods games.

Overall, I think that this paper provides an interesting discussion on individual learning in n-player games, highlighting the range of long-time behaviors that can be supported by a relatively simple cognitive model for a variety of asymmetric multiplayer games. The paper is well-written, and the description of the mathematical model is presented in an accessible way to reach a wide audience of researchers interested in evolution and learning in games coming from fields including applied mathematics, biology, economics, and physics. Due to both the interesting core results of the paper and the extensive potential for related future work, I think that this paper would be a good fit for Dynamic Games and Applications. However, I think that the authors could potentially improve the paper by discussing in greater depth connections between the current research and work in the learning literature on perturbed best-response dynamics (also called logit dynamics) [1–5] . With these considerations in mind, I believe that this paper will be appropriate for publication in Dynamic Games and Applications after incorporating a small to moderate level of revision to address connections to the literature on learning in games.

We thank the reviewer for their thorough assessment. We appreciate the positive comments and the relevant questions that were raised.

Major Comment

The primary comment I have on this manuscript regards the existing work in the literature on stationary distributions under introspection dynamics for multiplayer games [1–3]. In particular, both Auletta and coauthors and Marden and Shamma have studied convergence to stationary distributions in repeated n-player, asymmetric games under the logit learning rule used to model a smoothed best- response dynamics which allows individuals to pick new strategies with a probability proportional to the exponential weights introduced by Blume [6]. The main difference between the logit learning dynamic and introspection dynamics is that the logit rule requires agents to calculate the payoffs they would receive under each of their possible actions, rather than evaluating the payoff for a single randomly selected strategy. While the general of transition probabilities would seem to be similar under the two learning rules as both rules favor strategies that would improve individuals payoffs, but it is possible that the random selection of a strategy to imitate under introspection dynamics may result in different stochastic trajectories than under the logit dynamics in which all possible potential strategies are considered under any transition event.

At a minimum, I think that it would be a good idea to include more discussion and references to the literature on logit / perturbed best-response dynamics, explaining similarities and differences in that existing work to the behavior of introspection dynamics. There could also be interesting opportunities for future work in studying the differences between the stationary distributions produced by introspection or logit dynamics, or whether the convergence times to equilibrium or metastable states may be different for the two stochastic processes [1, 7]. In particular, for the case of coordination games, per- haps one of the processes is more successful in acheiveing a higher weight placed upon the equilibrium with higher payoff in the stationary distribution.

We thank the reviewer for raising these points. We agree that introspection dynamics and logit-response (smooth/perturbed best-response) have interesting connections that should be addressed. We now make some precise comparisons between the two processes. Particularly, we add to the discussion in the last section and add three new propositions comparing the two processes in the Appendices.

As the reviewer rightly says, introspection dynamics and logit-response are different (microscopic) update rules. In introspection players only need to compare payoffs between the counterfactual scenario and the current scenario to compute the probability of switching to a new strategy. Whereas, in logit response, players need to construct the entire probability mass function over their strategy space to sample their new strategy. This makes introspection a model with simpler cognitive assumptions. At the same time, the probability function with which players adopt new strategies in both processes has a similar shape – they adopt a new strategy (a_i) with a probability proportional to the exponential of the payoff scaled by selection ($e^{\beta(\pi_i(a_i,\mathbf{a}_{-i})}$), but the normalization constants differ. As such, it is natural to ask if and when the long-run behaviors (and times to convergence) resulting from these processes are equivalent.

To make this comparison, we take the logit dynamics with asynchronous learning (Blume 1993, Alós-Ferrer and Netzer 2010) where the single updating player is uniformly drawn from the pool of players. We find that the stationary distribution of the logit rule under these assumptions is equivalent to the one of introspection dynamics for three types of games: 2-strategy games, potential games and additive games. We provide proofs for these results in the new Appendix 2. It is intuitive that the processes match for 2-strategy games. When there are only 2 strategies, introspection dynamics enforces that the random alternative strategy is the (only) one not being currently played. As such, the comparison is among all strategies present in the strategy set, just like in logit-response. Hence, the transition matrices are the same. This is no longer true for more than two strategies. In general potential and additive games, the stationary distribution of the two processes are identical even though the transition matrices may be effectively different. We note that it is also interesting to look at the limit of strong selection. In this limiting regime, it seems that there are a number of games (that do not lie in these above three cases) whose stationary distributions match. Still, a possible general answer might not be trivial to obtain. For example, we find instances of coordination games whose stationary distributions match and others where they don't. We give some concrete examples here. We look at two coordination games where the payoffs when individuals play the same action are the same for the two players and are the same between the two games. The payoff matrices are

	a_1	a_2	a_3
a_1	$5.0, 5.0$	$2.8, 4.7$	$1.0, 3.0$
a_2	$3.5, 3.3$	$4.0, 4.0$	$2.8, 2.4$
a_3	$4.4, 2.2$	$3.9, 2.3$	$3.0, 3.0$

(1)

	a_1	a_2	a_3
a_1	$5.0, 5.0$	$3.8, 4.1$	$1.3, 3.2$
a_2	$4.7, 2.2$	$4.0, 4.0$	$2.6, 3.2$
a_3	$3.4, 1.2$	$2.7, 1.8$	$3.0, 3.0$

(2)

When selection strength is high, in the first game, logit dynamics selects the (a_1, a_1) equilibria (i.e. the game is at state (a_1, a_1) with probability close to 1), while introspection dynamics selects (a_3, a_3). In the second game, the opposite occurs. Thus, we cannot immediately conclude if one of the dynamics is better at selecting the higher payoff equilibrium in multi-equilibria coordination games.
Another interesting example where the two processes differ is the Rock-Paper-Scissors. If we consider the payoff matrix

	R	P	S
R	$0, 0$	$-1, 1$	$1, -1$
P	$1, -1$	$0, 0$	$-1, 1$
S	$-1, 1$	$1, -1$	$0, 0$

, (3)

we see that for a high selection strength, introspection dynamics predicts a uniform distribution over all outcomes, whereas logit dynamics predicts a uniform distribution over all outcomes except ties, i.e., the states (R, R), (P, P) and (S, S). The reason is the following. In introspection dynamics, at a given state where a player is loosing (getting payoff -1), they might face the opportunity to switch their strategy to a tie (getting payoff 0). This is a likely transition as *i*) with the same probability a player chooses one of the two alternative strategies, and *ii*) given the payoff difference is positive for the two options, they will very likely select it. However, in logit dynamics with sufficietly high selection, players are close to best-responders; thus, if a player is loosing, they will most likely choose the winning strategy rather than the tie strategy. The difference between the two processes is essentially that logit works as a (perturbed) best-response and introspection as a *better*-response.
Overall, we do not yet provide a complete relation between introspection dynamics and the logit-response (and other learning models). We believe such an extensive analysis, including mixing times, could be part of potential future work as it is interesting and relevant on its own. For the present paper, we decided to rather focus mainly on the extension of the previous results to multiplayer games.

Minor Comments

I also have the following minor comments that may be helpful for revising the paper. In the middle paragraph of page 17, I am a little confused by the claim that the introspection dynamics framework is more efficient than studying stochastic evolutionary game theory. These two frameworks are only equivalent when viewing both the stationary distribution of a learning process and the long-time behavior of stochastic evolutionary game theory as solution concepts for the underlying n-player game. However, one could also ask whether the stochastic dynamics of a population of individuals using a introspection-like update rule would result in a stationary distribution with similar weights on strategies as the stationary distribution achieved under learning in a repeated game. Similarly, perturbed best-response dynamics have been used to study both the dynamics of a large population in one-shot games [4, 5, 8] and for the long-term strategy frequencies achieved by learning in a repeated game [1–3].

We agree that it is an interesting question whether the resulting stationary distributions from the stochastic dynamics of a population (of introspective players) and from the learning dynamics of a repeated game are similar. We also hint at this in the last part of the last paragraph of the discussion section. We informally checked that for 2-player, 2-strategy games, the stationary distributions are qualitatively similar. But a systematic investigation is still undone. Perhaps again this could be part of future work. Also, we take the opportunity to point out that introspection dynamics was compared to a population model of pairwise imitation in the case of 2-player, 2-strategy symmetric games in Couto et al. 2022 (https://doi.org/10.1088/1367-2630/ac6f76). All in all, we think that we can look at the difference between a learning process and population dynamics in two ways. Fundamentally, they are very different. While one tries to describe the decision-making process at the group level, the other contains evolutionary considerations (there is a population where strategies can evolve, e.g. through pairwise imitation or a birth-death process). Depending on the context, one or the other framework might be more suitable for what the modeler intends to study. From a different angle, we might want to know how a particular game is played and what are the more likely outcomes. In that case, we indeed believe that the two frameworks (learning models and population models) can be used as alternative solution concepts for the underlying games. Perhaps they can even complement each other. Thus, we argue that introspection dynamics is more efficient than a stochastic EGT model in that sense.

At the bottom of page 21, there appears to be a typo in Equation (47). Should $\beta\pi_k(q_k, l_{-k})$ be replaced by $\beta\pi_k(q_k, \mathbf{p}_{-k})$?

In Equation (50) on page 22, I think it would be helpful to recall what quantity is being calculated on the lefthand side of the equation. It could be potentially confusing to break up a multiline equation with text explanations.

References

[1] V. Auletta, D. Ferraioli, F. Pasquale, P. Penna, and G. Persiano, "Convergence to equilibrium of logit dynamics for strategic games," in Proceedings of the Twenty-Third Annual ACM Symposium on Parallelism in Algorithms and Architectures, pp. 197–206, 2011.
[2] V. Auletta, D. Ferraioli, F. Pasquale, and G. Persiano, "Mixing time and stationary expected

social welfare of logit dynamics," Theory of Computing Systems, vol. 53, pp. 3–40, 2013.
[3] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation," Games and Economic Behavior, vol. 75, no. 2, pp. 788–808, 2012.
[4] W. H. Sandholm, "Pairwise comparison dynamics and evolutionary foundations for nash equilib- rium," Games, vol. 1, no. 1, pp. 3–17, 2009.
[5] W. H. Sandholm, Population Games and Evolutionary Dynamics. MIT Press, 2010.
[6] L. E. Blume, "The statistical mechanics of strategic interaction," Games and economic behavior, vol. 5, no. 3, pp. 387–424, 1993.
[7] V. Auletta, D. Ferraioli, F. Pasquale, and G. Persiano, "Metastability of logit dynamics for coor- dination games," in Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1006–1024, SIAM, 2012.
[8] J. Hofbauer, S. Sorin, and Y. Viossat, "Time average replicator and best-reply dynamics," Mathematics of Operations Research, vol. 34, no. 2, pp. 263–269, 2009.

Thank you for pointing these out. We now corrected the typo and the equation and add some of the suggested references.

Reviewer 2

The authors of this paper consider multiplayer games to capture the interactions in daily life involving more than two individuals. They incorporate introspection dynamics to study asymmetric multiplayer games where players compare their current strategy to an alternative strategy and decide whether to switch based on the payoff difference between these two strategies at each time step. The previous results of introspection dynamics for games with two players are extended to games with an arbitrary number of players. The authors derive analytic expressions for a few special cases including additive games and symmetric games with two strategies. And they obtain numerical results for more general cases.

Their work is mathematically solid and interesting. It is novel and adds to the literature of both introspection dynamics and multiplayer games. Moreover, it brings forth a deeper understanding of the complexities of real social dilemmas. And I do enjoy the set of symbolic notations and the proofs of propositions and corollaries in the reading process. Nevertheless, there are a few issues I would like to address.

We thank the reviewer for their nice summary and comments.

The authors compare their methodology with those in previous studies. I am curious to know whether it is possible to support it with concrete examples to ensure readers understand the significance and relevance of their work and how it can be translated into practical solutions (in particular, in Introduction and Discussion and Conclusion). They are encouraged to avoid leaving readers with the impression that it is a paper with technical jargon and complex language that may hinder comprehension.

We agree that the original manuscript lacks a thorough comparison to other models. We hope that with the new propositions and comments mentioned above (please see the reply to Reviewer 1), it becomes more precise how introspection dynamics compares to previous models, namely to the logit-response.

The figure captions can be more clear and more concise. I am not sure if it is a good idea to use "we" throughout.

Apart from these, it seems that periods are missing here and there at the end of an expression. To name a few, Equations (1) – (12).

Also, is it a "multiplayer game" or a "multiplayers game"? Please double check.

Above all, I would suggest acceptance of this manuscript with minor revisions.

Thank you for pointing these out. We now corrected the typos. It should be "multiplayer games".