First of all, we would like to thank the editor and the reviewers for their feedback. Their comments have been extremely helpful and constructive. In the meantime, we have revised the manuscript accordingly. The changes are substantial. In particular, we have addressed the following issues:

* As suggested by reviewer #1, we discuss the relative complexity of reactive-n strategies and of reactive-n counting strategies.
* Also suggested by reviewer #1, we explore the effect of errors, analytically and with simulations (see also our new **Figures Sx, Sy**)
* In the simulations of our original manuscript, there were some interesting patterns in the evolving strategies. Reviewer #1 suggested to explore these patterns in more detail. We do this in our revised manuscript (see also the new **Figures Sx, Sy, Sz**, and **Table Sw**).
* As suggested by reviewer #2, we discuss the prior literature on direct reciprocity with longer memory (see also our new **SI Section X**, in which we summarize several key papers).
* Motivated by reviewer #2, we now also explore the evolutionary dynamics of memory-*n* strategies and memory-*n* counting strategies (see our new **Figure Sx**).
* Again, motivated by reviewer #2, we compute the volume of the set of partner strategies, and we compare it to the volume of stable defecting strategies (see our **Figure Sy**). To this end, we have also added an extra subsection to our **SI** in which we characterize all stable defecting strategies among the reactive-2 and reactive-3 strategies. We compare the respective results to earlier work by Stewart & Plotkin (Scientific Reports, 2016).

As a result, we believe the manuscript has improved substantially. We would like to thank the reviewers for providing us with the relevant feedback. Please find a detailed response to their comments below.

Editor Remarks to Author:

Both reviewers find much to recommend in the paper, but both have important suggestions for a major revision, which must be carried out before another review.

**Reply:** Thank you for considering our paper, for providing positive feedback, and for encouraging us to submit a properly revised manuscript. The reviewers’ comments were extremely helpful in refining the paper. Over the last months, we have derived several new mathematical results, and we have run further simulations to address these comments. Please find all details in our response to the reviewers below.

Reviewer #1:

Suitable Quality?: Yes

Sufficient General Interest?: No

Conclusions Justified?: No

Clearly Written?: Yes

Procedures Described?: Yes

Supplemental Material Warranted?: Yes

Comments:

This paper introduces new formal results for reactive with longer memories based on a novel algorithm to identify partner strategies in the repeated prisoners dilemma. Reactive strategies are a less complex subset of automata strategies that only condition on the co-player's last moves rather than conditioning both players' moves (e.g. Tit-for-Tat is a reactive strategy but Win-Stay-Lose-Shift is not). Prior work has mostly focused on strategies with short memory, but it is possible in practice that this is insufficient to cover the kinds of strategies used by people. This paper aims to close this gap by creating an algorithm that lets them tractably analyze reactive strategies with longer memories.

The author's key insight is that for any opponent's reactive strategy, there is an associated self-reactive strategy that yields the same payoffs. They build on this insight to find tractable conditions for testing whether a reactive-n strategy is a partner strategy (cooperative equilibrium strategy). The number of conditions seems to grow exponentially with the size of the memory, but the authors claim these conditions boil down to reducing the cooperation rate proportional to the number of defections in the memory. Finally, the formalism is applied to reactive-n counting strategies, which are a subset of reactive-n strategies, and the formalism is tested empirically.

Overall, I think this is interesting work that may move the field forward. However, I have some significant concerns and reservations about the implications of these analyses and the conclusions that can be drawn, which currently discount the impact this work might have:

**Reply:** Thank you for carefully reading our manuscript, and for the constructive feedback below.

First, the authors make some strong claims about the value of memory when comparing reactive-n and reactive-n counting strategies. However, this comparison seems to be made by naively mapping the variable n as a notion of memory. The reactive-n counting strategy only need to represent a single number and increment / decrement that number based on whether they experience a defection or cooperation. To be more concrete, a reactive-256 strategy will require 256 bits of memory, while a reactive-256 counting strategy only requires 8 bits of memory. This is one of the key results in the paper, taking about 1/4 the length of the abstract - either this notion must be refined, or the paper does not support one of its key claims.

**Reply:** Thank you for this helpful feedback. It seems to us the reviewer raises two concerns.

First, the reviewer argues that when we compare reactive-*n* and reactive-*n* counting strategies, the variable *n* might not properly reflect the true memory requirements of a strategy. After all, to represent a reactive-256 strategy, a player indeed needs to memorize the exact sequence of the co-player’s past 256 decisions (for example, the sequence CCDCCC…DCD, a string with 256 bits). On the other hand, a reactive-256 counting strategy only reacts to how often the co-player cooperated during the last 256 rounds. This is a number between 0 and 256. If we represent this number in binary notation, this information only takes 8 bits. (Perhaps more precisely, it takes 9 bits because there are 257 integers between 0 and 256; but obviously that does not diminish the argument in any way). So this observation suggests that the memory requirements for a reactive-*n* counting strategy are far less stringent than for a reactive-*n* strategy.

We think this is an extremely interesting observation. Yet it turns out that the issue of a counting strategy’s complexity is even more subtle. To see why, consider a player with a reactive-256 counting strategy. Moreover, suppose this player holds the number 157 in memory (i.e., the co-player cooperated 157 times during the last 256 interactions). That information will be sufficient for the player to choose the correct cooperation probability in that round. However, that information alone will not allow the focal player to correctly update its memory state. To see why, suppose the co-player cooperates in that round. As a result, the focal player might either have to update its memory to 157 or to 158. Which of those two options is correct depends on the co-player’s behavior 256 rounds ago (the piece of memory that is no longer relevant after the current round). If the co-player cooperated 256 rounds ago, the new state of the memory is 157. Otherwise, it is 158. Hence, a player with a reactive-256 counting strategy still needs to remember what exactly happened 256 rounds ago. By a similar argument, it follows that the player generally needs to have a record of the co-player’s actions during all of the last 256 rounds. We conclude that with respect to the number of bits required to implement a given strategy, reactive-*n* counting strategies are as complex as reactive-*n* strategies.

As the second concern, the reviewer is worried that we make strong claims that might not be backed up by our results. Here, the reviewer is particularly pointing to our abstract. We appreciate this feedback, because we certainly do not wish to misrepresent the nature of our results. Unfortunately, however, we had difficulties identifying those parts in our abstract that might be misleading. In our honest opinion, the abstract just describes how we define certain strategies, what we do with them, and what we find. If the reviewer has more specific advice on how we should revise our abstract, we would be extremely grateful.

**Changes:** Regarding the first comment, we have added a new paragraph to our discussion, in order to better explain the memory requirements of the different strategy spaces. This paragraph follows the arguments given above. Regarding the second comment, we carefully revised the entire article (not just the abstract), to avoid formulations that we felt could be misleading.

Second, a key motivation for longer memory strategies (according to the authors) is the need to study longer memory strategies. They write, "Longer memory seems particularly relevant for noisy games, where people occasionally defect because of unintended errors". However, there is no discussion of whether this framework applies to the repeated prisoner's dilemma when there are errors. Do reactive-n strategies that have longer memory actually perform better under higher error rates? This should be investigated analytically and empirically since this is a key motivation for the relevance of this class of strategies.

**Reply:** That is right, in our introduction we briefly reviewed the experimental literature on repeated games. There, we mentioned that participants seem to be particularly likely to adopt higher-memory strategies when decisions are subject to errors.

While this literature provides some motivation for our study, our original submission did not consider the effect of errors specifically. This had two reasons. The first is simplicity: Analyzing the set of partner strategies in conventional repeated games is already a considerable mathematical challenge. Doing the same analysis for repeated games with errors adds another layer of complexity. In fact, even for the comparably well-studied case of memory-1 strategies, it is still an open problem to describe the partner strategies for games with errors.

The second reason is that the existing empirical literature is just one motivation for our study. Perhaps our main motivation is a theoretical one: How far can we push our understanding of direct reciprocity when we drop the usual – but somewhat extreme – assumption of one-round memory? Can we develop mathematical tools that allow us to make progress? A better understanding of these questions can also help us better understand existing results. For example, one of the most prominent examples of a strategy with one-round memory is Generous Tit-for-Tat (GTFT). This strategy cooperates with certainty if the co-player cooperated in the previous round. But even if the co-player defected, GTFT cooperates with a certain probability. In the donation game, this probability is 1-*c*/*b* (where *c* and *b* are the cost and benefit of cooperation). Our analysis of reactive-2 and reactive-3 strategies shows that this upper bound on GTFT’s cooperation probability is just one special case of a more general pattern. We feel these insights are important in their own right, independent of the empirical literature we mentioned.

Having said that, we fully agree with the reviewer that further results for games with errors would be desirable. As suggested by the reviewer, we have therefore explored these games both numerically and analytically. Our simulations suggest that our earlier results are quite robust. We still find that higher memory has a positive impact on cooperation when players use reactive strategies. We still find that higher memory fails to have this positive effect when players only use counting strategies (see our new **Figure Sx**). In addition, our new analytical results suggest that almost all of the previously described partner strategies remain approximate partner strategies for sufficiently small error rates (we describe these results in more detail in the **SI**). Moreover, we also describe a class of strategies that are Nash equilibria for all error rates.

**Changes:** In the revised main text, we have added a short subsection on errors at the end of our Results section. The relevant details are described in the Supporting Information (see **SI Section X.Y**, and the new **Figures Sx** and **Sy**). In addition, we have also revised our introduction section. In particular, we explain that the mentioned empirical evidence is one motivation for our study, but perhaps not the key motivation.

Third, another conclusion from this work is that knowing the sequence of moves matters, not just the count. But it seems from looking at the conditions, to be a partner strategy, the order does not matter in a straightforward way (equation 1). For reactive-2 strategies, all that matters is the sum of p\_dc and p\_cd with equal weight. Likewise, for many of the conditions in reactive-3 strategies in equation 2.

**Reply:** That’s right. In terms of evolution, our simulations suggest that it is important to remember the exact sequence of moves. Yet, in all of the analytical conditions for reactive-2 partner strategies, only the co-player’s average cooperation rate matters (across the last two rounds). For an interpretation of this mismatch, it is important to note that our analytical results only characterize a strategy’s static stability. In contrast, in evolutionary simulations, it is also important how likely such strategies are to emerge in the first place, and how easily they can be invaded in a stochastic evolutionary process.

Moreover, it is worth to point out that our definition of partners strategies requires . That is, such players cooperate with absolute certainty if their co-player was fully cooperative in the past. However, stochastic evolutionary processes such as ours do not generate strategies in which this condition is exactly met. Instead, the entries of a new mutant strategy are generated by drawing a random number between [0,1]. As a result, mutant strategies may well have a cooperation probability close to one, but this probability will not be equal to one. This observation is important, because for strategies with , the entries and are no longer interchangeable. We come back to this observation in more detail in our response to the reviewer’s fourth comment below.

**Changes:** We now mention these arguments at the end of our subsection on our evolutionary simulations.

Finally, looking closely at the conditions in equation 2, it does not seem to generalize reactive-2 strategies where the cooperation rate needed to decrease by c/2b - is there some other pattern that applies, or is that intuitive explanation only valid for reactive-2?

**Reply:** For reactive-2 partner strategies, one of our conditions in Eq. [1] is indeed

The general pattern here is the following: the proportion of D’s in the considered histories on the left hand side is 1/2 (There are two strings ‘CD’ and ‘DC’, with four letters overall; out of these four letters, one half of them are ‘D’). Therefore, also the relevant factor on the right-hand side is 1/2. The same pattern also applies to all our conditions for memory-3 strategies, see Eq. [2]. For example, one of these conditions is:

Here, the proportion of ‘D’s on the left-hand side is 2/3. Hence, the maximum cooperation probability on the right-hand side is correspondingly reduced by . So again, players reduce their cooperation probability in proportion to how often the co-player defected in the past.

Of course, one may interpret the reactive-2 strategies as a subset of the reactive-3 strategies. More precisely, reactive-2 strategies are those reactive-3 strategies for which for all . In particular, we should be able to recover the conditions for reactive-2 strategies within the reactive-3 strategies. This is indeed possible. For example, taking the condition and plugging it into the second inequality in Eq. [2], we obtain

and

These two inequalities exactly reproduce the middle inequality in Eq. [1] for reactive-2 strategies within the reactive-3 strategies. Again, the cooperation rate is reduced by *c*/(2*b*).

**Changes:** We now explain this general pattern more clearly when we describe Eq. [2] in our main text.

Fourth, when looking at the simulated results, the abundance of reactive-2 strategies seems to differ substantially between p\_cd and p\_dc, but this is not predicted by theory (and neither is p\_cd < p\_dd). What explains this result? I was also hoping to learn more about what it is about the timing of cooperation that makes it important. Tit-for-two-tats and other known versions make some intuitive sense, but can these intuitions be mapped onto the conditions and the framework presented here?

**Reply:** This is an excellent observation. Indeed, as mentioned earlier, our analytical results on reactive-2 partners seem to suggest that the two variables and are interchangeable. For example, consider the two reactive-2 strategies . Then, depending on the parameters of the game, our conditions suggest that either both of these strategies are partners, or none of them is. Yet when we look at **Fig. 4A**, our evolutionary simulations seem to favor strategies with . That is, our evolutionary process seems to favor strategies that tend to punish a co-player’s defection immediately, rather than with one round delay. In our earlier submission, we did not provide any explanation for this asymmetry.

To provide such an explanation, we first checked whether this asymmetry arises consistently. In our earlier submission, **Fig. 4A** was based on ten independent simulations. When we increase the sample size to twenty, we still find the asymmetry (we have updated **Fig. 4A** correspondingly).

To get a better sense of this asymmetry, we looked at two specific examples, the strategies and mentioned above. Using simulations, we verified that the two strategies are equivalent with respect to their evolutionary properties. For example, we considered a scenario in which a homogeneous resident population adopts strategy (or , respectively). We then used simulations to explore how much time it takes on average until the resident strategy is successfully invaded by a mutant strategy. We get the same result for and , see our new **Figure Sx**.

However, as noted in one of our previous responses, our evolutionary process does not generate boundary strategies where one of the cooperation probabilities is exactly one. Rather this process might generate slightly perturbed variants like and . Once we compare these two strategies, we observe a difference. For example, we find that it takes more time to invade a homogeneous population with strategy , compared to one with strategy . This asymmetry arises because payoffs in a homogeneous -population are slightly higher than in a homogeneous -population (see our new **Figure Sy**). This payoff difference arises because two -players are on average better able to recover full cooperation once one of the players defected. We describe these new results in detail in our new **Figure Sz** and **Table Sw**. These findings suggest that for a reciprocal strategy’s performance, it is better to punish defection without delay.

Overall, these results suggest that strategies like and are indeed equivalent. However, this equivalence no longer holds when these two strategies are slightly perturbed. In that case, we find that the first strategy is more resistant. This explains why on average, our simulations produce strategies with .

**Changes:** We now discuss this asymmetry in our evolutionary results in the main text, when we present **Fig. 4**. In addition, we provide further details in the **SI** (in particular, we provide more data in **Figures Sx-Sz** and **Table Sw**). Thank you for making us explore this asymmetry in more detail!

Reviewer #2:

Suitable Quality?: Yes

Sufficient General Interest?: Yes

Conclusions Justified?: Yes

Clearly Written?: Yes

Procedures Described?: Yes

Supplemental Material Warranted?: Yes

Comments on Significance Statement:

Results are limited to reactive strategies, but conclusions are stated far more broadly, especially in the significance statement.

**Reply:** Thank you for sharing this feedback. For most of our original significance statement, we merely introduce important concepts, such as the notion of ‘nice strategies’ and ‘partner strategies’. When we describe our conclusions, we do mention that we characterize these strategy sets among the ‘longer memory reactive strategies’. However, we agree that this sentence only mentions ‘reactive strategies’ in passing; we should have mentioned them more prominently. We have now rephrased the significance statement accordingly. We also realized that the last two sentences of our significance statement might come across as too general. We have rephrased them too.

Comments:

The authors characterize the Nash equilibria that sustain cooperation among the space of reactive memory-n strategies, for a two-player 2x2 iterated game. The analysis results in explicit conditions for n=2 and n=3, and also for arborary n when restricted to "counting" strategies that merely count how many times an opponent has cooperated. The authors are also interested in strategy evolution in a population, and they compare their analytical results on cooperative memory-n strategies to monte-carlo simulations in the limit of weak mutation.

Finding or characterizing Nash equilibria for any sizeable space of strategies in an iterated game is a hard problem. The authors have made real progress on such a problem, when restricted to "reactive" strategies that condition only on your opponent's plays. This is very solid work, and it is made possible by extending the approach of Press & Dyson, Akin 2012, and also Park.

Perhaps the most intuitive and elegant result has to do with the requirements for a reactive counting strategy to be Nash and ensure full cooperation in the donation game. Such a strategy must reduce the chance of cooperating by a constant factor (c/nb), for every co-player's defection observed across the prior n-rounds. This is more than simply a characterization of Nash equilibria for reactive counting strategies, but also a very nice intuition for how cooperative Nash equilibria must respond to instances of defection in the opponent.

**Reply:** Thank you for this summary of our work, and for the positive evaluation!

As the authors are surely aware, finding Nash equilibria does not always predict the outcome of evolution in population -- hence the notions of ESS and ESS\_N in finite populations. In this part of their paper, though, the authors resort to simulations to determine what reactive strategies dominate in populations undergoing payoff-biased imitation. (As it happens, at least among reactive strategies, that Nash conditions do a fairly good job a predicting the types of strategies that are evolutionary robust.)

**Reply:** Indeed – even if a strategy is a Nash equilibrium, it does not need to be evolutionarily stable. Similarly, such a strategy does not need to emerge in evolutionary simulations. Still, we find the notion of a Nash equilibrium to be useful, even in the context of evolutionary dynamics. For large and well-mixed populations, we can only expect a strategy to be resistant to mutant invasions if it is a Nash equilibrium. And as the reviewer mentions, our analytical conditions on partner strategies are in quite good agreement with the strategies that appear in our simulations.

My main critique has to do with how the authors treat prior literature on this topic of long-memory strategies in iterated games. Part of the problem here is that they mis-represent prior work as being purely based on simulation, when in fact much prior work is strictly analytical and in even greater generality than the reactive strategies studied here. The other part of the problem is that the authors have not compared their qualitative results observed in evolutionary simulations to the analytical results in prior literature for how memory length effects the volume of cooperative evolutionary stable strategies.

I am supportive of publication after the authors give a complete account of prior work and make effort (see specific suggestions below) to contextualize their results (especially the results that seem at odds with prior studies of long-memory strategies).

**Reply:** Thank you for the feedback. The above critique is well taken. After re-reading our manuscript, we agree that our description of the previous literature was somewhat superficial. Similarly, we did not sufficiently connect our simulation results to the results of previous papers. We thus appreciate that the reviewer brought up these topics, and we are happy to address them. For more details, please see our responses below.

Critique 1) Regarding prior work on long-memory strategies, the authors state that "previous studies considered simulations for small n (56-59), or they analyzed the properties of a few selected higher-memory 134 strategies (60-62)." This is not an accurate description of prior work.

In particular, McAvoy & Nowak (2019) [not cited in this sentence] contains very detailed analytical work on a fascinating class of learning strategies, that is strict super-set of memory-1 strategies. McAvoy shows, eg, that such learning strategies can punish a defecting opponent over multiple rounds; and that such strategies have much greater power to shape the region of feasible payoffs compared to simple memory-1 strategies (which results in convex hulls). The authors are clearly aware of this prior work on long-memory strategies, and so they should discuss it and not dismiss it as purely simulation, when it is detailed analysis.

Likewise, the results of Ueda 2021 (memory-2 strategies) are not discussed or compared to the authors' own results.

Likewise, the authors describe Stewart et al (Ref 57) as based purely on simulation, which is false. Stewart et al derives analytical conditions for the space of memory-n strategies that result in either pure cooperation or pure defection, and that resist selective invasion by any mutant strategy in a populations of size N. (The authors are clearly aware that Ref 57 is based on mathematical analysis, because they cite Stewart as having previously developed a generalization of ZD strategies for memory-n, in their supplement.)

In general, the authors should discuss prior analytical work on long-memory strategies for iterated games, especially in evolving populations, and compare those prior results to their own results. There are some striking differences (see below), which must be reconciled and discussed.

**Reply:** We fully agree. We should have discussed the prior literature in more detail, to avoid any ambiguity about our contribution. We do this now.

In short, the papers by McAvoy and Nowak (2019) and by Ueda (2021) describe some properties of certain higher-memory strategy sets. McAvoy and Nowak consider ‘reactive learning strategies’; Ueda considers ‘zero-determinant strategies’. Both of these papers prove a number of beautiful results. However, neither of them explores the set of Nash equilibria within the respective strategy sets. In particular, they do not describe partner strategies, as we do in our article.

Perhaps the article closest to ours is the one by Stewart and Plotkin (Scientific Reports 2016). For the most general version of their model, they consider a more comprehensive setup than ours: they consider memory-*n* strategies for games among *m* players. Using an elegant coordinate transformation, they describe the set of all evolutionary robust strategies that can sustain cooperation. This set is closely related to our notion of partner strategies. Importantly, however, their conditions (their Eq. [2]) only give an implicit characterization of this set. More specifically, their inequalities include terms that become increasingly hard to compute for larger values of *m* or *n* (even numerically). In addition, these terms depend on the specific mutant strategy considered. In principle, one would still have to verify those conditions for uncountably many mutant strategies. In their SI, Stewart and Plotkin argue that in practice, only four extremal values of need to be tested. However, the authors do not elaborate on the necessary procedures in more detail. Moreover, they do not derive any conditions that are independent of the terms .

In contrast, we prove that for a general reactive-*n* strategy to be a partner, one only needs to test all ‘mutants’ that adopt a pure self-reactive strategy. We use our characterization to give explicit conditions for reactive-2, reactive-3 and reactive-*n* counting strategies. We show that in general, testing only four specific mutant strategies is not enough. For example, for reactive-*n* counting strategies, it takes *n* independent conditions to verify whether a given strategy is a partner (see our Eq. [3]).

**Changes:** In the revised main text, we provide a more detailed description of the state of the literature. In addition, we have added an extra section to our **SI**, in which we summarize the most relevant articles (including the articles mentioned above). In particular, we also describe how those articles differ from our study.

Critique 2 (which is related): Figure 4 seems to report evolutionary simulations only for reactive memory-n strategies. But I am not sure if the main results of the figure (notably the distinction between panels C and D) will hold beyond reactive strategies. I suggest the authors repeat the exploration in Figure 4 with the full space of memory-2 or memory-3 strategies.

**Reply:** Sure! We have now repeated the respective simulations for memory-1, memory-2 strategies, as well as for memory-1, memory-2, and memory-3 counting strategies. Qualitatively, we obtain the same results as in our **Figure 4**. Again, we find that higher memory can promote the evolution of cooperation. But again, we find that for higher memory to be advantageous, strategies need to take into account how often each player cooperated. For counting strategies, the beneficial effect of higher memory is notably reduced.

**Changes:** We now briefly mention our new results on memory-*n* strategies in the main text. In the **SI**, we provide more details (see also **Figure Sx**).

This is an important question, I feel, because the authors use the results of Figure 4 to conclude that the order of events must be remembered, not just the number of cooperative events, for a longer memory to benefit cooperation. But this general conclusion is not justified beyond the (restrictive) space of reactive strategies. In fact, Ref 57 has shown (analytically, and verified by simulation) that the opposite trends holds when counting strategies are not just reactive: among the space of strategies that track how many times you cooperated and your opponent cooperated, the volume of evolutionary robust strategies that induce full cooperation is made larger, relative to volume of robust defecting strategies, as memory length increases. In other words, strategies that count how many times both you and your opponent cooperated (which is a very natural thing for humans to count!) are very different than simply counting how many times your opponent cooperated. The authors should discuss this distinction, and they must certainly temper their conclusions about the need to keep track of order of events, which is limited to reactive strategies alone.

**Reply:** Thank you, this is a very important observation. Indeed, in the paper by Stewart & Plotkin, the authors compute the volume of the set of all evolutionarily robust memory-*n* strategies that sustain mutual cooperation. They compare this set to the set of evolutionary robust strategies that lead to mutual defection. They find that as they increase *n*, the volume of cooperative strategies increases, relative to the volume of defecting strategies.

In our revised manuscript, we now do the same analysis for our setup. Just as Stewart & Plotkin we find that the volume of partners is increasing in *n*, relative to the volume of defecting strategies. This result is true for both reactive-*n* strategies, and for reactive-*n* counting strategies. Hence, with respect to the size of the respective strategy sets, there is no difference between Stewart & Plotkin’s results on memory-*n* strategies and our results on reactive-*n* strategies.

However, these results only establish that as we increase *n*, partner strategies make up a larger portion of the strategy space, relative to the defecting strategies. These results do not directly imply that evolutionary processes would also increasingly favor the evolution of partner strategies (and hence the evolution of cooperation). Indeed, our evolutionary simulations for reactive-*n* and for memory-*n* strategies suggest that this is only true when players take into account the exact order of events. For counting strategies, we find no advantage of higher memory when it comes to the evolution of cooperation.

Unfortunately, it is difficult to compare these evolutionary results to Stewart and Plotkin’s. We study the evolution of cooperation when a player’s memory is kept fixed. In Stewart and Plotkin’s evolutionary simulations (their Fig. 4), a player’s memory is allowed to co-evolve. Moreover, when Stewart and Plotkin report simulation results, they do not report the resulting average cooperation rates, as we do. Instead, they report the ‘Ensemble mean strategy frequency’ of cooperators and defectors (their Fig. 4a,b). We assume these curves represent the respective volumes of cooperative and defecting strategies, given the population’s mean memory. Because they do not report how often players cooperate on average, we cannot tell whether they find any positive effect of memory on evolving cooperation rates.

**Changes:** We fully agree with the reviewer that we need to discuss our evolutionary results in light of the work of Stewart & Plotkin. We do this now, both in the main text and the SI. In particular, we describe that just as in Stewart & Plotkin, the relative volume of partner strategies is increasing in the players’ memory capacity (relative to the volume of defecting strategies). This is true both for reactive-*n* and for reactive-*n* counting strategies. To show this result, we have considerably extended our **SI**. In addition to the partner strategies, we now also provide an explicit characterization of the stable defecting strategies.

With respect to our evolutionary simulations, we still find that higher memory only has a positive effect for reactive-*n* (and memory-*n*) strategies, but not for the respective counting strategies. However, as suggested by the reviewer, we now interpret these results more carefully. In particular, we only draw conclusions for those strategy spaces that we explicitly explored.