Artificial Life Manuscript Submission

The Dynamics of Social Interaction among Evolved Model Agents

Haily Merritt [1,2], Gabriel J. Severino [1], Eduardo J. Izquierdo [1,2]

Corresponding: Eduardo J. Izquierdo (edizquie@iu.edu)

1. Cognitive Science Program, Indiana University Bloomington
2. Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington

Abstract. We revisit the perceptual crossing simulation studies, which are aimed at challenging methodological individualism in the analysis of social cognition by studying multi-agent real-time interactions. We offer three advances: First, we evolve and test agents in rigorous conditions to build confidence in their ability to solve the task in a more human-like way. Next, we transform the sensor from discrete to continuous, which facilitates an in-depth dynamical analysis of how agents respond to the different objects in the environment. Finally, we examine agents' behavior with other agents to determine how they perform with a partner whose neural controller is different from their own. Altogether, our findings emphasize the opportunities for dialogue between artificial and human perceptual crossing studies and highlight the contributions of simulation studies for understanding social interactions.

Keywords: evolutionary simulations, continuous-time recurrent neural networks, simulated social interaction, perceptual crossing experiments

1 Introduction

Recent years have seen an increasing number of scholars call for studies of social cognition to consider the *interaction* between social agents as the entity of interest, as opposed to the behavior or cognitive and neural processes of one social agent (De Jaegher et al., 2010; Pfeiffer et al., 2013; Schilbach et al., 2013). Concurrently with this theoretical push, tools and perspectives from complex systems science (e.g. dynamical systems theory and network theory) have facilitated a richer understanding of the role of interactions per se for social agents since they make no assumptions of methodological individualism (Froese & Di Paolo, 2010; Froese & Fuchs, 2012). An interactionist approach coupled with complex systems science tools is well-poised to elucidate the mechanisms supporting social interactions since the *interaction*–as opposed to the *interactor*s–is the object of study (De Jaegher et al., 2010; Froese & Di Paolo, 2011).

One paradigm in which such advances have been clear for the study of social interaction is the perceptual crossing paradigm (Auvray et al., 2009). In these experiments, participants are placed in a simple one-dimensional virtual environment and tasked with identifying when they believe they are interacting with another participant. However, they are not aware of the true identity of their partner. The nature of these tasks is such that they cannot be solved by one participant alone, highlighting the importance of mutual interaction and joint recognition. The ability for the nuance of social interaction to be studied in a simplified minimal setting helps to remove unnecessary complexities of an experiment, while also opening up valuable collaborations with researchers in artificial life. This has set forth a series of fruitful dialogues between both empirical and theoretical efforts to understand perceptual crossing (Di Paolo et al., 2008; Froese & Di Paolo, 2008, 2009, 2010; Rohde & Paolo, 2008), and by extension social interaction in general (Di Paolo, 2000; Di Paolo et al., 2008; Froese & Di Paolo, 2008; Iizuka & Di Paolo, 2007; Iizuka & Ikegami, 2004; Ikegami & Iizuka, 2007; Quinn, 2001; Reséndiz-Benhumea & Froese, 2020; Reséndiz-Benhumea et al., 2021; Williams et al., 2008).

Specifically, simulations studies have shown themselves to be quite invaluable, with the ability to manipulate factors that otherwise would be impossible in human subjects, like making the objects infinitely small or rigorously unpacking the role of the task setup using psychophysical and other analyses (Froese & Di Paolo, 2009, 2010). Importantly, while we may not be able to have a fully nuanced understanding of internal dynamics underlying human behavior in perceptual crossing experiments, simulation studies offer us a unique opportunity to understand this connection through the mathematical analysis of dynamical systems theory.

Paramount to studying interactions *per se* is disavowing hindering assumptions made implicitly in methodologically individualistic research. This includes taking seriously the diversity of social agents, valuing contextual differences, among other things. Social psychologists have advocated vigorously for such approaches (Cikara et al., 2022; Lewis Jr, 2022). In artificial life research, and specifically within the perceptual crossing paradigm, endorsing such perspectives might involve moving more thoroughly from uniform virtual populations or clones to diverse agents. Much of the existing work on perceptual crossing examines the behavior of a pair of clones. Assuming clonality is useful for making the paradigm more tractable to rigorous analysis and has indeed yielded a rich repertoire of results (Di Paolo et al., 2008; Froese & Di Paolo, 2009, 2010). As of yet, it is unclear whether clones are required for successful perceptual crossing or whether distinct agents can also solve the task.

Recently, we addressed several open questions about the simulation of perceptual crossing by systematically examining and reporting on the conditions that do and do not lead to successful crossing (Izquierdo et al., 2022). First, simulation studies have all relied on the introduction of a sensory delay for the agents to perform the perceptual crossing task successfully (Di Paolo et al., 2008; Froese & Di Paolo, 2009, 2010). Crucially, the practical need for delays in the models has been considered a potentially important component for the explanation of the adaptive performance of the task in human participants and has motivated psychological studies. However, the necessity of a sensory delay in human participants is unlikely (Iizuka et al., 2015). We showed that a sensory delay is not necessary for human-like behavior of the artificial agents that align with previously reported work (Izquierdo et al., 2022). Removing the sensory delay yielded two patterns of behavior: agents that crossed only a handful of times and agents that crossed perpetually. Because only perpetual crossers had been reported in previous literature (Di Paolo et al., 2008; Froese & Di Paolo, 2009, 2010), we assumed the perpetual crossing strategy was preferred because the agents necessarily continuously interact. It was unclear whether minimal crossers had been evolved before, since most existing work only reported one solution, instead of an ensemble of successful solutions. In our third experiment, we evolved agents that cross perpetually by modifying the fitness function to select for both the proximity of agents as well as a high number of crossings. Additionally, we implemented two strategies to make the paradigm more rigorous and systematic. First, we excluded initial transient dynamics from our fitness function, which improved our precision in measuring the percentage of trials where agents found each other. Second, instead of using a stochastic fitness function, which varied the starting locations of the agents, we used a deterministic fitness function that systematically tested a large swath of starting locations (Izquierdo et al., 2022).

We identified several areas of opportunity that provide additional rigor, systematicity, and intrigue to the perceptual crossing task. First, from observation and psychophysical analysis, we determined that many agents relied on the fixed object or shadows to successfully engage in crossing. This represented a deviation from human behavior, where the fixed object and shadows are distractors. In other words, relying on the fixed object or shadow to achieve perpetual crossing could be seen as cheating on the task. To address this, we evolve agents across a set of conditions in which the fixed object or shadow is not always

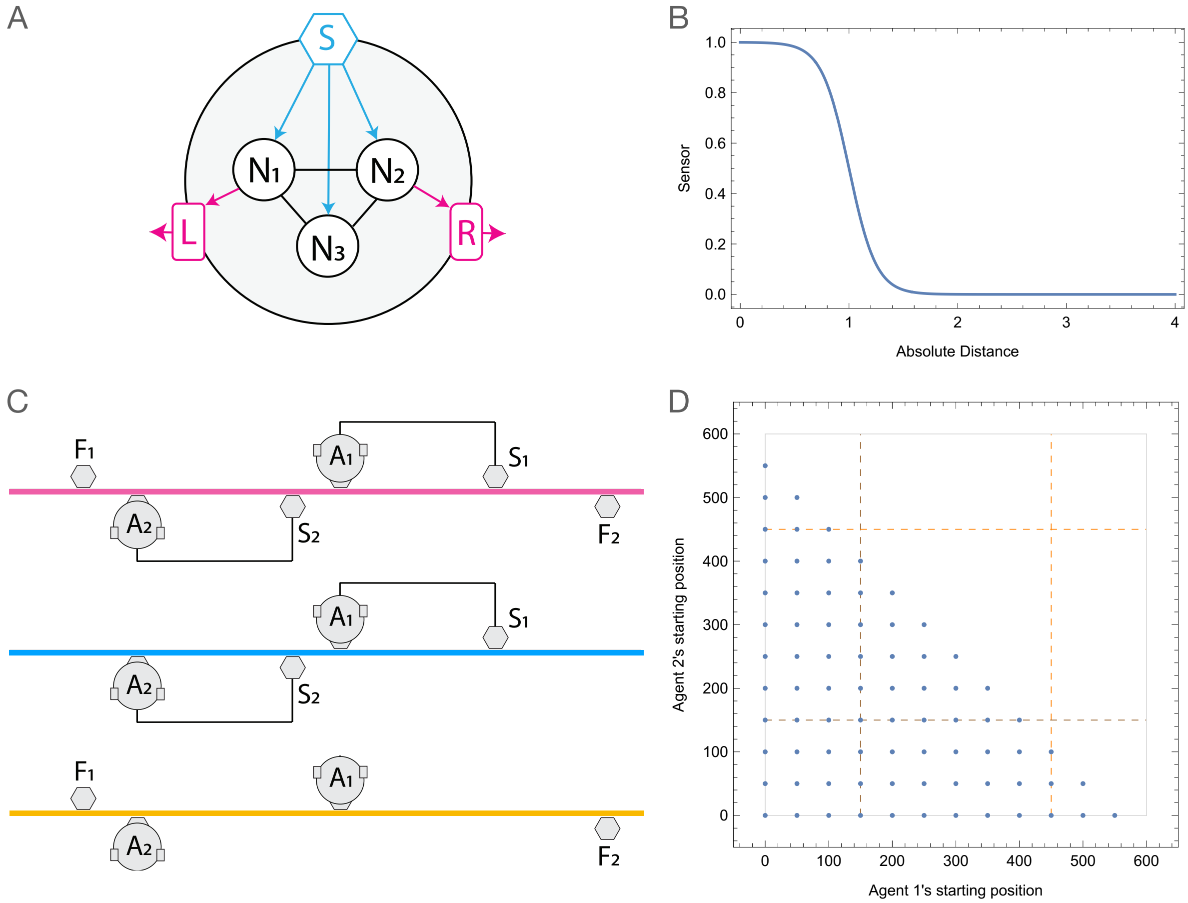

Figure 1: Task and agent setup. (**A**) Each agent has a sensor (cyan) that can send information to all N neurons (black). The neurons in the circuit are fully interconnected, including self-connections (not depicted). The output from one neuron drives the left motor and another neuron drives the right motor (magenta). The neural circuits in the two agents are identical (i.e., they have the same parameters) except where otherwise noted. (**B**) Instead of a discrete (on/off) sensor, we introduced a continuous sigmoidal sensor that is more active the closer the agent is to another entity. (**C**) The task takes place in a 1-dimensional ring where two agents face each other. In the Original condition (magenta), each agent can sense the other's avatar (A), a shadow of the other's avatar (S), and a fixed object (F). We introduced two additional conditions during evolution. In the No-Fixed object condition (cyan), agents can sense each other and each other's avatars, but there is no fixed object. In the No-Shadow condition (yellow), agents can sense each other and the fixed object, but the shadows have been removed. These additional conditions ensure that evolved agents do not rely on the fixed objects or shadows to solve the perceptual crossing task. (**D**) The fitness evaluation for each circuit involved 78 different starting conditions (blue points), obtained by systematically varying the starting position of the two agents around the 1D ring (600 units of space) in steps of 50 units of space, excluding conditions that are symmetrical (where the agents start in the opposite starting conditions, but otherwise identical).

present. Second, because the agent pairs were always clones, it is unclear if any of the solutions were robust enough to detect other agents more generally. That is, such simulated social interactions would be more compelling if the agents interact with agents *other than themselves*. In addition, the discrete nature of the sensors made the dynamical analysis difficult. Given the centrality of dynamical analysis to the interpretation of simulated perceptual crossing, an easier approach to studying the model's dynamics seems valuable. As such, we redefined the sensor to have a continuous sigmoidal function.

In this paper, we extend the work on perceptual crossing to address the open questions above. The rest of this paper is organized as follows. In the next section, we describe the perceptual crossing task and the setup of the agents for all experiments. We had a three-part approach to our analyses. First, we discuss the evolutionary and behavioral results of successful perceptual crossers. Next, we perform a thorough dynamical analysis of successful robust circuits. Third, we explore how well robust circuits can perform with agents whose neural controllers are different from their own. Finally, we conclude with a general discussion of the experimental results and outline some directions for future work.

2 Methods

Our experimental design is largely identical to Experiment 3 in our previous paper (Izquierdo et al., 2022): no temporal delay, same one-dimensional ring environment and agent morphology (Fig. 1A), same circuit sizes (2 - 4 neurons), same CTRNN neural controller (Fig. 1B), same fitness function with a conditional component, no transient dynamics included in the calculation of fitness, and a fully deterministic experimental setup. For additional details on the neural controllers, experimental design, or fitness function, we point the reader to the Supplementary Materials section and our previous study (Izquierdo et al., 2022). Briefly, we achieved the fully deterministic setup by testing agents at each of 78 starting positions uniformly spread across the 1D ring (Fig. 1D). The conditional component of the fitness function was designed such that when an agent's fitness is above 0.99, agents were given additional points for perpetual crossing proportional to the number of crossings.

We note that in the present paper, we introduced three important changes to the experimental design: First, we changed the sensor from binary to continuous to facilitate dynamical analyses. Second, we reduced the timestep of integration from 0.1 to 0.05 during evolution. Part of the reason for both of these changes (the continuous sensor and the smaller timestep of integration) was to ensure that the neural circuits that evolution produced were actually fit to solve the task. In preliminary analyses, we noticed that most of the agents trained on the larger timestep of integration typically failed to solve the task when tested on a finer timestep. This problem was further aggravated when the discrete sensor was taken into consideration. Finally, we trained all agents on three different environmental conditions: (1) Original setup with all three objects–agents, shadows, and fixed objects–present, (2) No fixed object but including agents and shadows, and (3) No shadows but including agents and fixed objects. We introduced this final change, in the form of additional task conditions, to ensure that the agents did not rely on either the fixed objects or the shadows to identify each other. In preliminary analyses, we observed that although we could find successful agents that did not rely on either the shadow or the fixed object, this was mostly a matter of luck: A good proportion of successful agents under the original condition did indeed rely on those other components and did so in interesting ways. Altogether, given the 78 starting locations and 3 environmental conditions, each agent underwent 234 fitness evaluations, each lasting 800 units of time. The population consisted of 96 individuals. Each evolutionary run was performed for 1000 generations. We performed 100 evolutionary runs for each circuit size. Each evolutionary run was provided with a different seed.

3 Results

We present our results as follows: Section 3.1 describes the evolutionary and behavioral results of our robustness testing. This includes all changes we made to the perceptual crossing task to make it more rigorous and systematic. Section 3.2 presents our dynamical analyses of the neural circuits. We focus only on successful robust circuits. Section 3.3 describes our tests of the extent to which successful robust agents are really social. Here we perform analyses of pairs of successful agents to determine how well they can identify others who are different from themselves as well as analyses of successful robust agents with decoys. This approach marks a departure from previous simulation studies that analyzed only pairs of clones.

3.1 Evolution and behavior of robust perceptual crossing agents

In our previous study, we noticed that solutions sometimes relied on the fixed object or the shadow to solve the task. For example, one of the two agents would encounter the other's shadow, which would prepare it to engage the next stimulus, 'knowing' it would be the agent. By including conditions where these other objects are sometimes not present, we are deliberately making the task more rigorously focused on the goal: detecting mutual interactions. Thus, our first question was: Can we evolve agents who are successful at solving this more rigorous perceptual crossing task? We performed 100 evolutionary runs with two-, three-, and four-neuron circuits (Fig. 2). Across all circuit sizes, we found solutions that solved the problem nearly perfectly (>0.99). Consistent with our previous study, the number of successful solutions found increased with the number of neurons in the circuit: 3% of all two-neuron circuits, 12% of three-neuron circuits, and 19% of four-neuron circuits.

In order to focus our analysis only on solutions that are as robust as possible, our next step was to test the performance of all solutions even more thoroughly (Fig. 3). Specifically, we

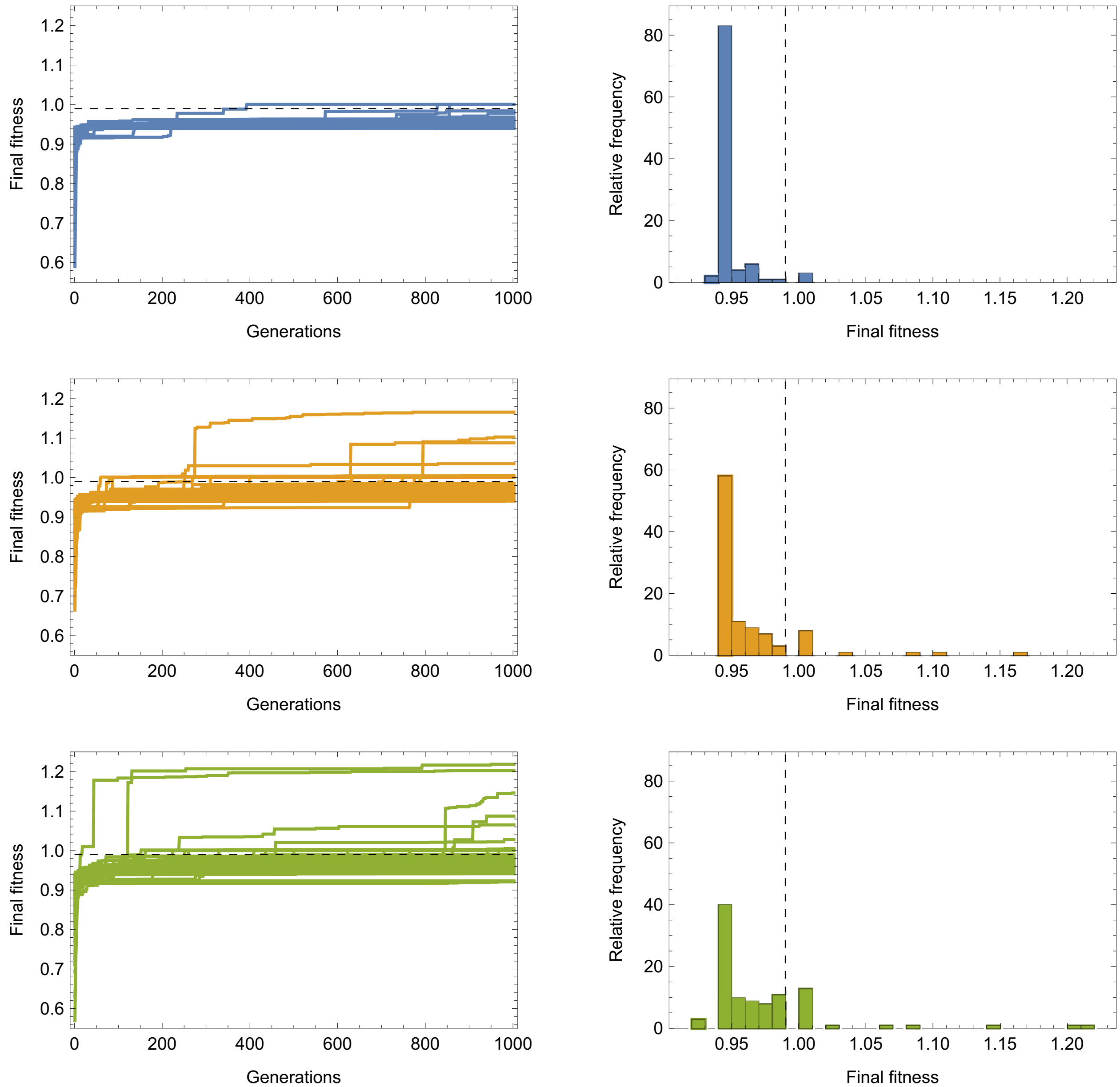

Figure 2: Evolutionary performance. We performed 100 evolutionary runs for two-neuron circuits (blue), three-neuron circuits (orange), and four-neuron circuits (green), shown top to bottom, respectively. In the left column, we show the fitness of the best individual in each of the populations as a function of generations. In the right column, we show the performance histograms for the final best solutions across those same conditions. The dashed line represents a performance of 0.99, above which the fitness function is modified to include the number of crossing as part of the measure of success. Some portion of the evolutionary runs across all circuit sizes produced solutions that surpassed the threshold: 3 two-neuron circuits; 12 three-neuron circuits; and 19 four-neuron circuits.

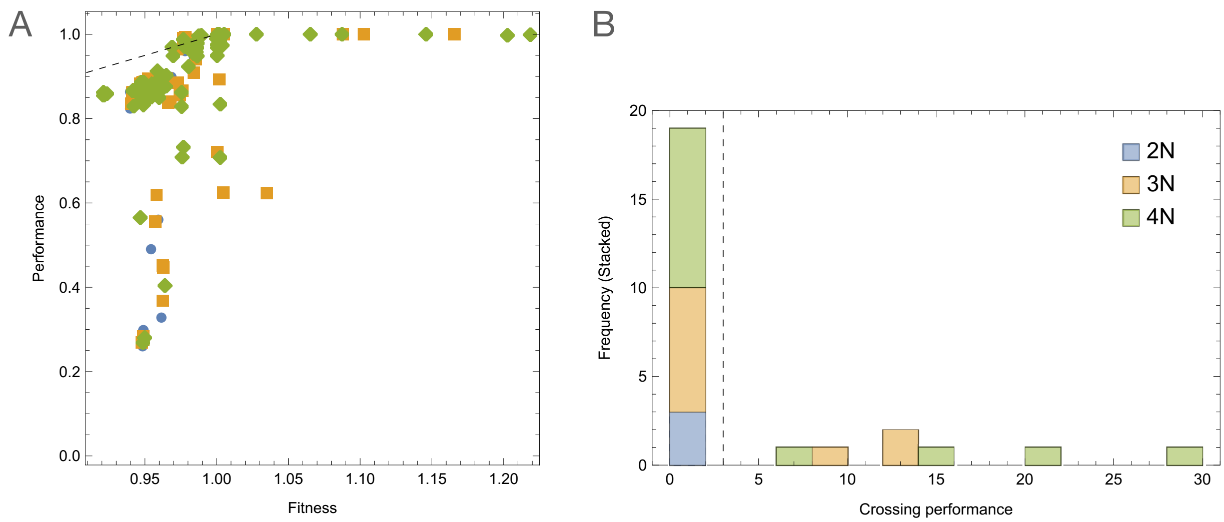

Figure 3: Ensemble of robust circuits. (**A**) Relationship between the final fitness of a solution evaluated during the evolutionary run and the more thorough evaluation of performance performed afterward. We further examined the performance of all 100 solutions on all three circuit sizes: two-neuron (blue), three-neuron (orange), and four-neuron (blue). The dashed line represents the region where solutions have the same performance and fitness. Solutions below the dashed line are those that have lower performance than their fitness. Solutions on or above that line are those whose performance on the more exhaustive evaluation matched or exceeded their original fitness. Note that fitness can be greater than one as well when the solutions have a large number of crossings; whereas the performance evaluation is agnostic in relation to crossings. We consider robust the circuits that obtained a performance > 0.99. (**B**) Number of crossing between interacting agents for the ensemble of robust circuits tested over a limited duration of time. Out of the robust circuits, we observed two different kinds of solutions. The dashed line represents the separation between these two groups. Some of the robust solutions cross only a couple of times; while a few of them cross perpetually.

reduced the time-step of integration from 0.05 to 0.01, increased the number of starting conditions from 78 to 36,000, and doubled the duration of the trial to 1600 units of time. We evaluated the performance of all 300 solutions (across the three different circuit sizes). Of these, 26 circuits demonstrated a highly robust performance (> 0.99) on this finer examination: 3 two-neuron circuits; 10 three-neuron circuits; and 13 four-neuron circuits. We refer to this ensemble of solutions as the 'robust circuits.' Note that this is a smaller subset of the original 34 circuits with a final fitness above the 0.99 threshold. As expected, some of those circuits did not perform well under the more thorough examination (Fig. 3A).

Further analysis of these robust circuits revealed that they could generalize well across a number of relevant unseen conditions, including changing the exact location of the fixed objects, and the relative location of the shadows.

As we first pointed out in our original study (Izquierdo et al., 2022), there are two different strategies used by the successful circuits to solve this mutual interaction task. The first group of circuits cross the other agent a limited number of times and then stop crossing; we call this group *minimal crossers*. The other group of circuits cross continuously as part of their strategy; we call this group *perpetual crossers*. How many of these successful robust circuits are perpetual crossers? In order to distinguish perpetual from minimal crossers, we characterized the average number of crossings over the last 100 units of time during the same performance evaluation described previously (Fig. 3B). Given the long initial transient, agents that exhibit any number of crossings are categorized as perpetual crossers; agents that exhibit zero crossings (but maintain close proximity to each other) are labeled minimal crossers. We observed that all 3 robust 2-neuron circuits are minimal crossers. Of the 10 robust 3-neuron circuits, 3 of them are perpetual crossers and 7 of them are minimal crossers. Of the 13 robust 4-neuron circuits, 4 of them are perpetual crossers and 8 of them are minimal crossers.

In order to understand how these circuits solve this mutual interaction task, we start by visualizing the pattern of behavior for one successful robust perpetual crossing circuit. We sampled the traces of the movement of the best-performing three-neuron circuit for one arbitrary starting condition (Fig. 4). At first, we observe that the two agents start 200 units of space apart from each other (Fig. 4A). Recall that both circuits are identical and they are flipped on either side of the 1-dimensional ring, so in the absence of a specific differentiating stimulus, they exhibit similar but opposite behavior: one turns clockwise and the other anti-clockwise around the ring. One of the agents encounters the fixed object (depicted as a dashed horizontal line in position 150). Panel (i) depicts the interaction

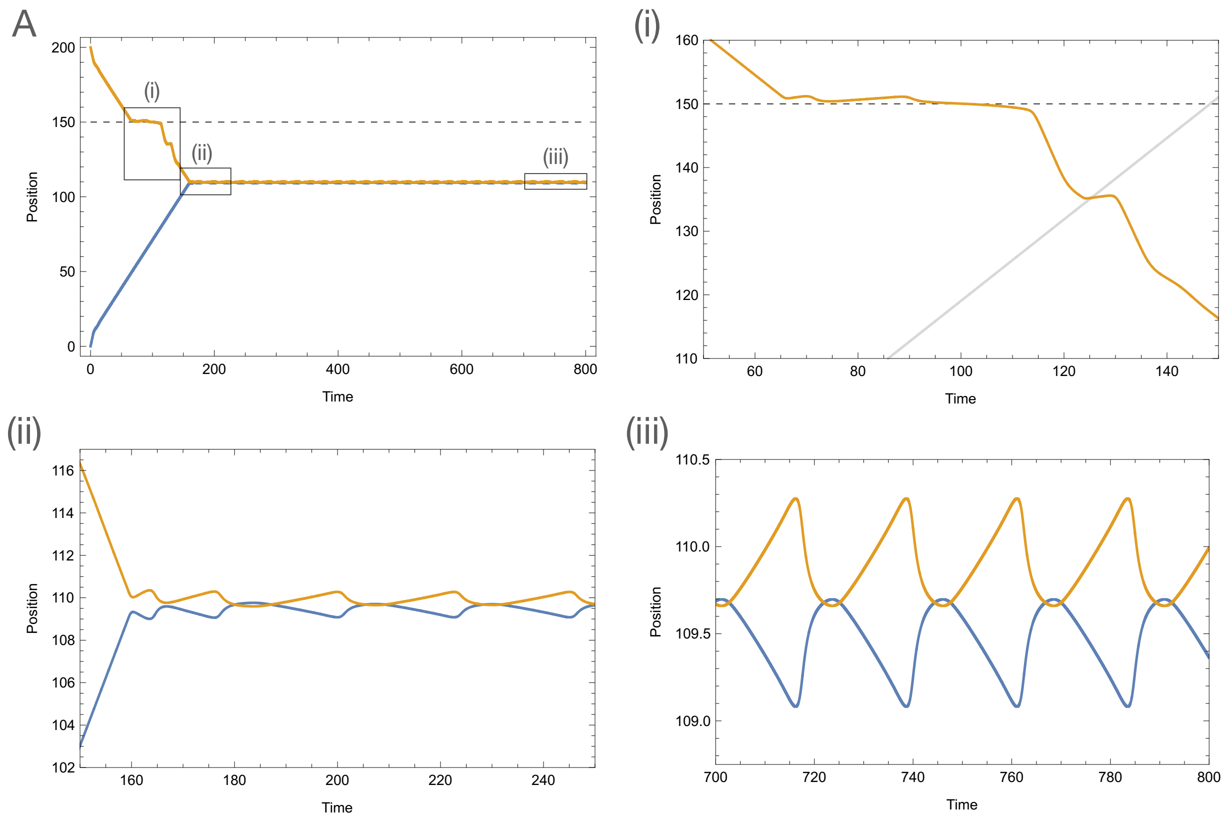

Figure 4: Behavior of a successful circuit. (**A**) Traces of two agents (blue and orange) from an arbitrary starting condition (one of the 36,000 tested during the robustness performance) for 800 units of time. The two agents are clones of each other. As can be seen, they successfully find each other, distinguishing between themselves and the other objects in the environment. We select three sections of the trace to zoom into as a way to consider the different kinds of interactions more closely: (**i**) Interaction between the agent and the fixed object (horizontal dashed line), and between the agent and the other agent's shadow (light gray trace); (**ii**) the beginning of the interaction with the other agent after interacting with its shadow; and (**iii**) perpetual crossing with the other agent.

between this agent and that fixed object. The agent senses it and responds by turning back towards the fixed object. This pattern of behavior occurs three times before the agent eventually moves past the fixed object, effectively deciding it's not the other agent. In that same panel, we see that shortly after the encounter with the fixed object, the agent encounters the shadow of the other agent (depicted as a light gray trace). Here again, the agent reduces its speed until it eventually turns around, moving towards the shadow. As the shadow keeps moving away from the agent, however, the agent quickly reduces its speed again and resumes its original direction of movement, realizing the object with which it was interacting was not the other agent. Fewer interactions are necessary for the agent to move past the shadow than to move past the fixed object. Panel (ii) depicts the start of the interaction with the other agent, shortly after the interaction with its shadow. The interaction looks relatively similar to that with the fixed object: the agent slows down, turns to move away, and then back towards the agent, and this process is repeated three times again, as with the static object, but this time the two agents engage in a repeating crossing pattern. Panel (iii) depicts the two agents interacting in this perpetual crossing behavior later in the trial. This behavior was typical across different starting conditions.

In short, the behavioral strategy used by successful robust perpetual crossers is to slow down upon detecting an object in the environment and then move back and forth over it. If the object also moves back and forth, the agent 'knows' it has found the other agent.

3.2 Dynamical analysis of the neural basis of behavior

Evolutionary results from the first section suggest the perceptual crossing task can be reliably and robustly solved by relatively small neural circuits. Furthermore, the behavioral analysis provides us with some insights into the different strategies the solutions used to achieve good performance. In order to better understand *how* these behavioral strategies came to be, we have to peek "under the hood" of behavior and take a look at the neural dynamics of these circuits and how the dynamics are shaped by interactions with the environment, following previous studies (Froese, 2018; Froese & Di Paolo, 2010; Froese & Fuchs, 2012). We proceed by analyzing the dynamics of one solution in detail. For continuity, we select the same agent whose behavior we characterized in the previous section (Fig. 4).

In order to understand the operation of one of these agents, it is particularly useful to isolate the nervous system first and study its autonomous dynamics as a function of the

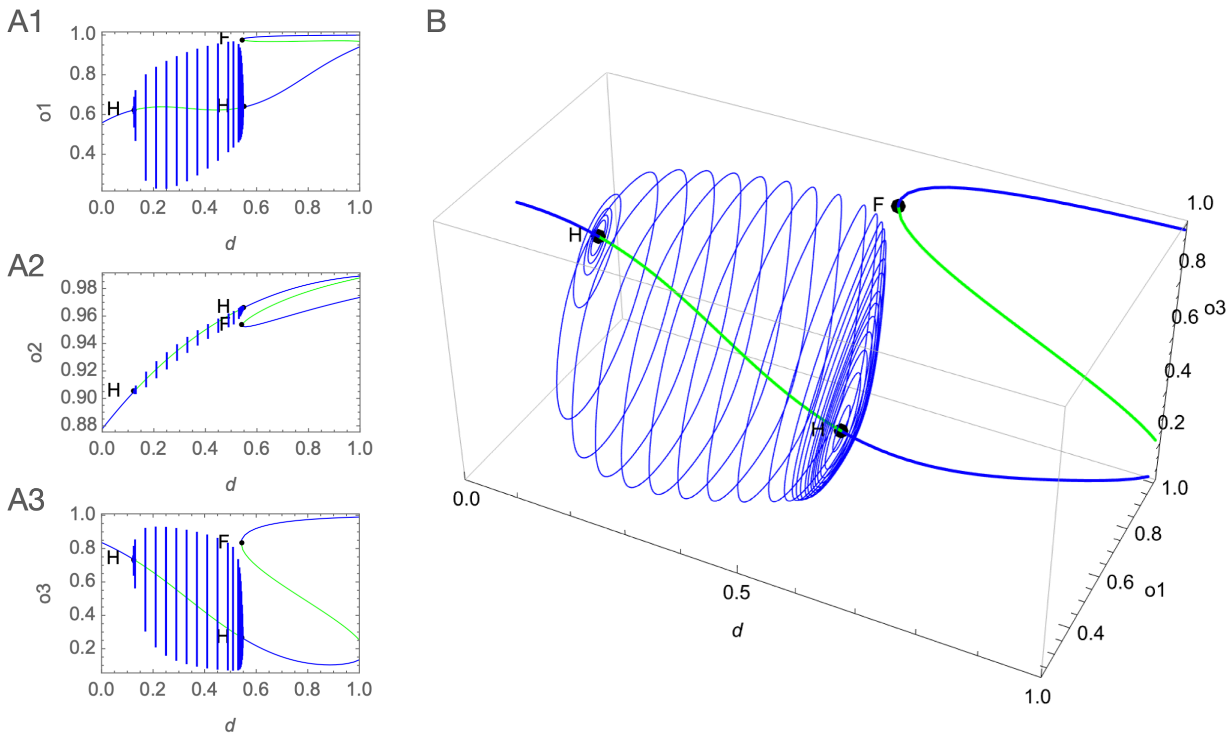

Figure 5: Example bifurcation diagram for one robust three-neuron circuit. (**A**) On the left side, we show one-dimensional slices through this four-dimensional bifurcation diagram, one for each of the three neurons (labeled accordingly). The x-axis (labeled d) represents the sensory value, from 0 (sensor off) to 1 (sensor on). Blue lines indicate stable equilibrium points for a given value of the sensor in our system. Green lines indicate saddle points. (**B**) On the right side, we show a two-dimensional slice of this space (for neurons 1 and 3). This circuit undergoes a supercritical Hopf bifurcation (labeled as H) when the sensor = 0.12. This can be characterized as a local bifurcation in which our stable equilibrium point loses stability, while a pair of complex conjugate eigenvalues cross the complex plane. This then arises in a stable limit cycle occurring around a saddle point. The limit cycle continues until the sensor = 0.54, in which the limit cycle terminates alongside a fold bifurcation into two stable equilibrium points separated by a single saddle point.

different values the sensor can take. By treating the sensor value as a parameter, we can perform a bifurcation analysis as the sensor transitions from a fully off (0) to fully on (1) (Fig. 5). When the sensor is fully off (sensor = 0), the neural circuit has a single stable equilibrium point, which we know experimentally corresponds to movement in one direction. When the sensor is fully on (sensor = 1), the neural circuit is bistable, as represented by the two stable points separated by a saddle point. As we will see, however, only one of those is ever approximated during ongoing behavior, which corresponds to movement in the opposite direction.

Because we introduced a continuous sensor, novel to the analysis of perceptual crossing agents is the dynamical transition that occurs as the sensor turns on. In this agent, some of the most interesting neural dynamics occur between sensor values of 0.12 and 0.58. As the agent begins to encounter an object, it receives a small perturbation to its sensor and undergoes a Hopf bifurcation; there is a change in the stability of its equilibrium point and the appearance of a stable limit cycle (periodic orbit). This limit cycle persists until the sensor value is approximately 0.58, at which point the system undergoes a fold bifurcation. This brings up a natural question: Is this limit cycle a deciding mechanism for distinguishing between an interacting agent and a non-mutually-interacting object?

To relate the dynamical traces in our bifurcation analysis to behavior, we transformed the four-dimensional space (i.e., sensor value and outputs of neurons 1, 2, and 3) into two dimensions. In doing so we make two key simplifications. The first simplification is to collapse the output of neuron 1 and neuron 2 into one dimension because the movement of the agent is determined by the difference between these two neurons (i.e., whether the agent moves clockwise or counterclockwise around the 1-dimensional ring depends on whether the output of neuron 1 is greater than the output of neuron 2, and vice versa). This collapse is particularly important because it allows us to directly relate neural dynamics to behavior. The second simplification is to collapse the sensory dimension. We achieve this by color-coding the different limit sets according to the sensor value for which they exist.

In order to understand how the non-autonomous dynamics of the neural circuit relate to behavior, we study this 2D representation of an agent's dynamics as it interacts with the other agents and objects in its environment (Figure 6. The stable attractor when the sensor is off captures the agent's movement at a speed of around 0.4 in one direction. As soon as a stimulus is sensed, the speed is reduced until the movement turns in the opposite direction. In the case of the shadow (Figure 6F), as the stimulus quickly disappears, the

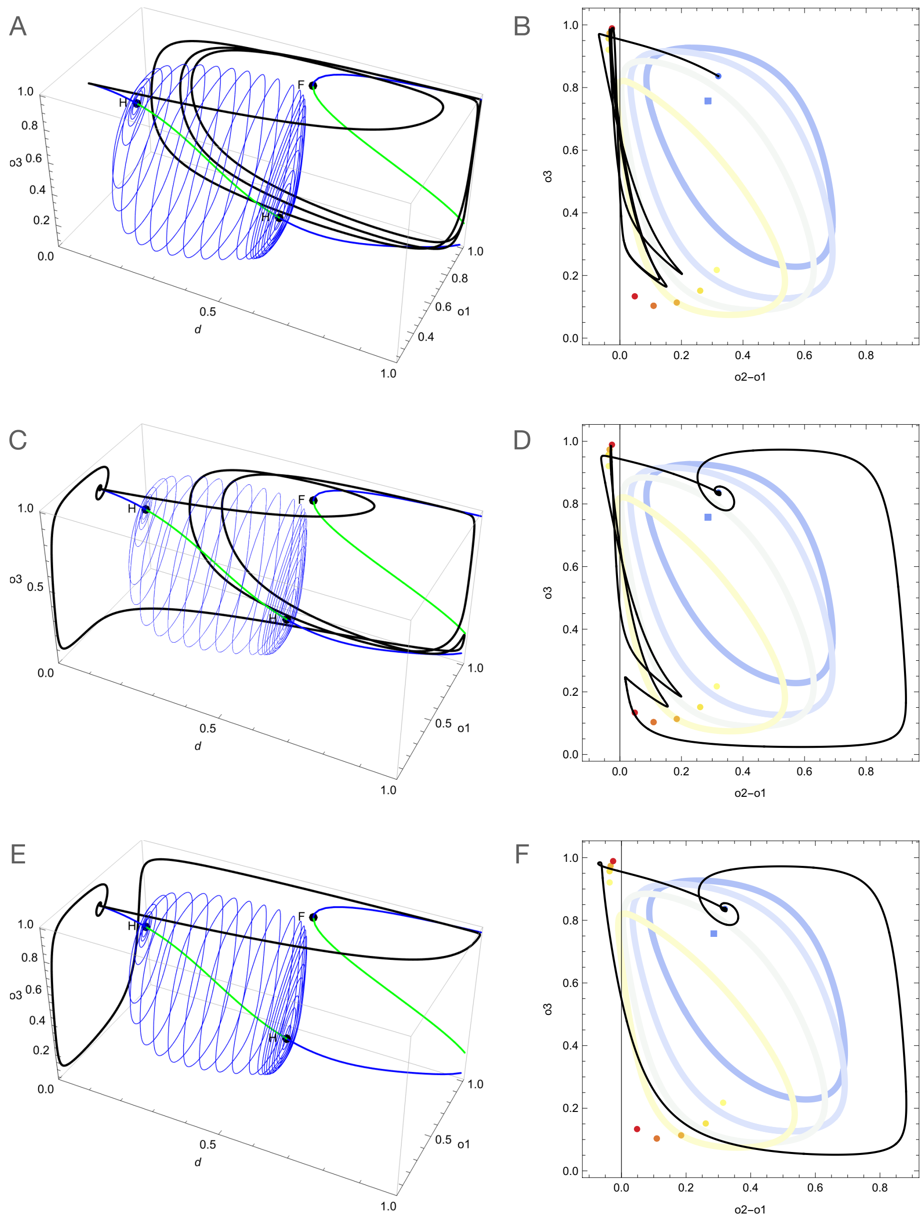

Figure 6: Different kinds of interactions. State trace of the three-neuron perpetual crosser as it interacts with another agent (top panels), a fixed object (middle panels), and a shadow (bottom panels). Panels on the left side, depict the traces in the 3-dimensional slice of the bifurcation diagram shown in Figure 5, overlaid on top of that same corresponding bifurcation diagram. Panels on the right-hand side, depict a lower dimensional transformation of that same space, where the outputs of neurons 1 and 2 are combined in a way that is relevant for behavior. Also, the limit sets of the system at different sensory values are depicted in different colors, from blue for no stimulus (sensor = 0) to red for full stimulus (sensor = 1). The traces of the state of the system as it interacts with each of the different other objects is shown in black.

movement resumes in the original direction relatively quickly. The interactions with the other agent and the shadow are more similar. In both, the sensor's activation shifts the movement in the opposite direction, which causes the sensor to turn off again, which drives the movement in the opposite direction again and causes the sensor to turn on again, and so on. The crucial difference between the fixed object and the other agent is that, in the case of the other agent, this pattern is maintained indefinitely (Figure 6B). In the case of the fixed object, however, the pattern is eventually disrupted (Figure 6D). Crucially, we can observe that in this agent, the limit cycle does not play a crucial role in the decision-making process.

The presence of a limit cycle within the dynamics of the neural circuit was intriguing enough that we decided to analyze all 26 of the robust circuits. Of the 3 robust 2-neuron circuits, all of which were minimal crossers, two had limit cycles, and 1 did not. Of the 10 robust 3-neuron circuits, there were 3 perpetual crossers and 7 minimal crossers, 2 of each containing limit cycles. Of the 13 robust 4-neuron circuits, 5 were perpetual and 8 minimal; only 2 perpetual crossers contained limit cycles. In summary, 50% of perpetual crossers contained limit cycles, while only 22.2% of minimal crossers contained limit cycles. In addition, across all circuit sizes, the presence of a limit cycle when an agent was tested against its clone did not predict whether it was a perpetual crosser ($\chi^2(1) = 0.914, p = 0.339$). Thus, while limit cycles are present in a good portion of the successful ensemble of robust circuits, they are not necessary for the success of the perceptual crossing task. Instead, what drives successful behavior is the switching between the different directions of movement as the sensor is activated.

3.3 How social are successful robust circuits really?

Thus far, all pairs of agents analyzed in the present and previous studies have been clones of each other. This feature raises the question: How social are successful robust circuits

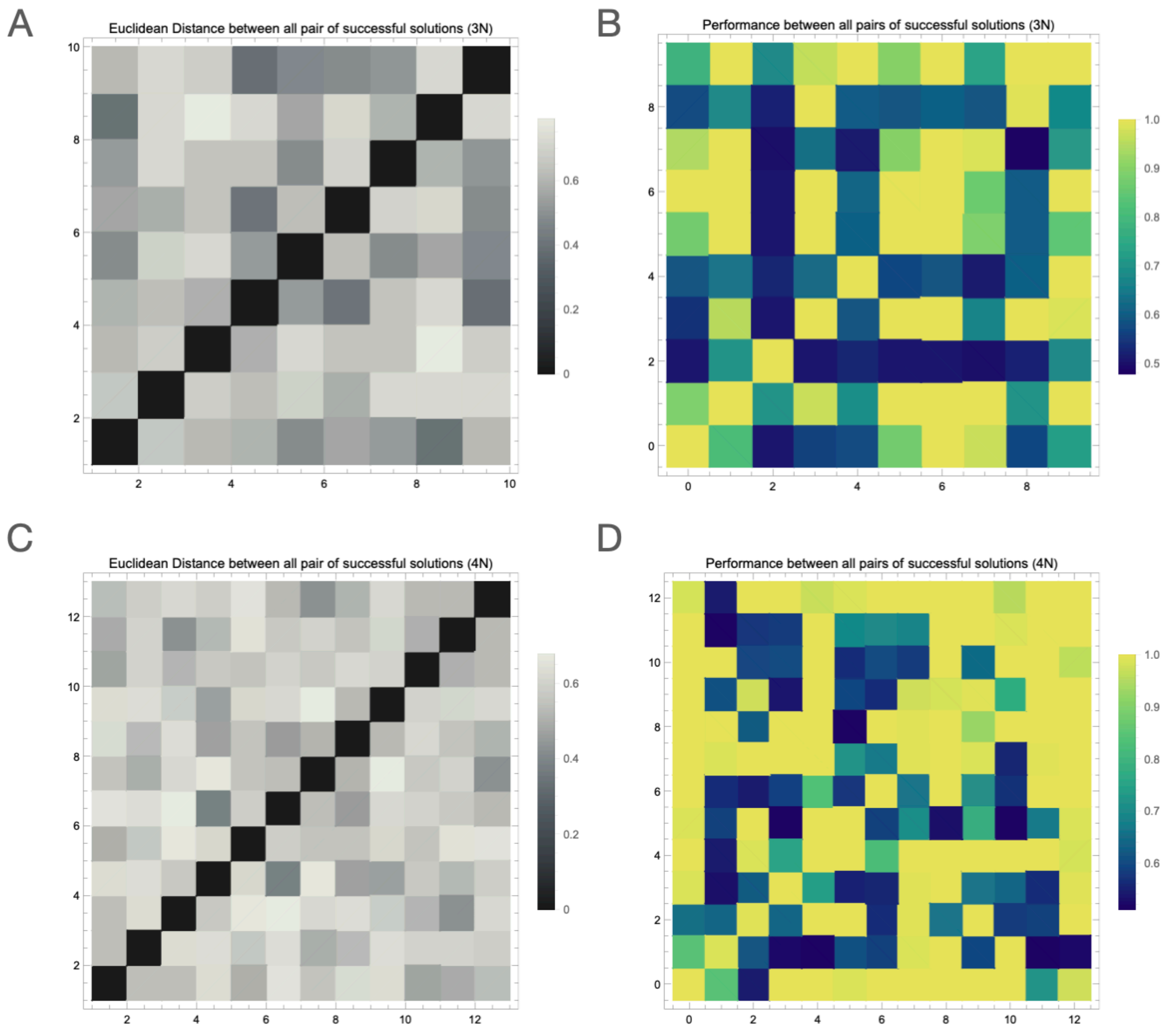

Figure 7: Inter-agent analyses. The top row shows all pairs of 3-neuron agents, and the bottom row shows all pairs of 4-neuron agents. Panels (**A**) and (**C**) show the Euclidean distance (normalized by the dimensionality of the parameter space) in parameter space between all pairs of 3-neuron and 4-neuron agents, respectively. This measure captures how similar a pair of agents' neural controllers are to each other. Panels (**B**) and (**D**) show the performance between all pairs of 3-neuron and 4-neuron agents. respectively.

really? Can a successful robust agent recognize another agent that is not its clone? To address this, we tested every pair of 3-neuron circuits ($N = 9$; Figure 7B) and every pair of 4-neuron circuits ($N = 12$) (Figure 7D). For this evaluation of performance, 0.5 is effectively random behavior or two agents that cannot mutually detect each other above chance, and 1.0 entails two agents that reliably find each other. Somewhat surprisingly, some pairs

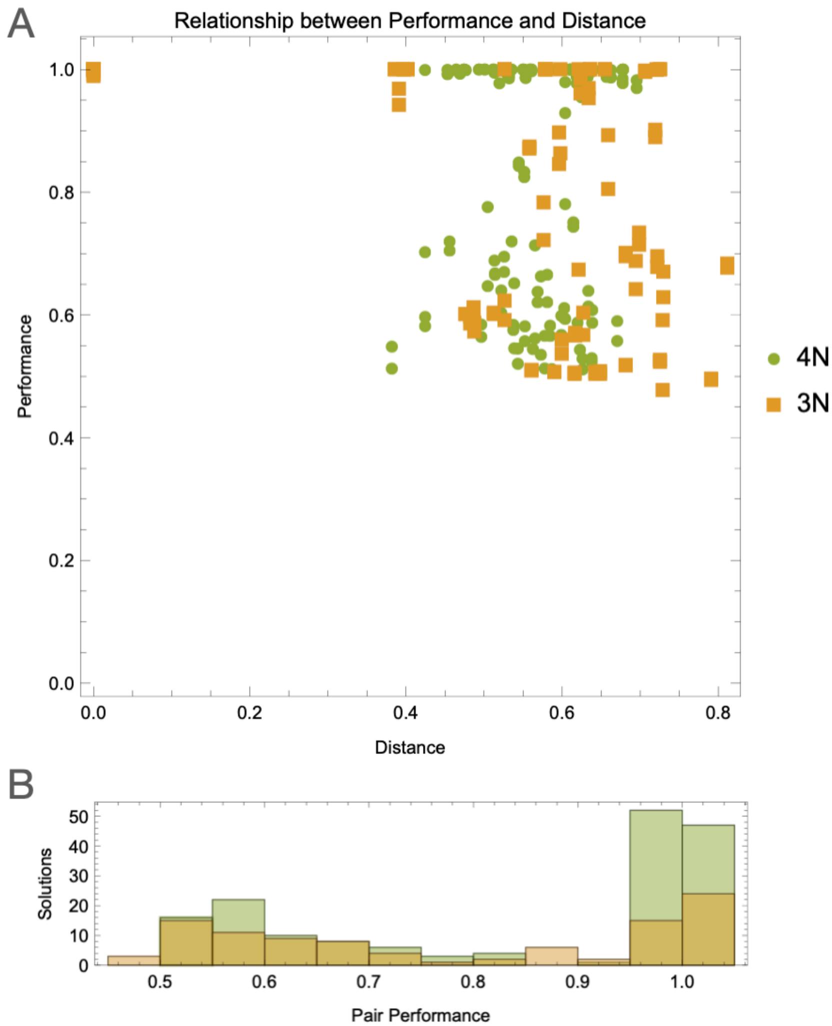

Figure 8: Relationship between performance and similarity. (**A**) Relationship between the ability of two different agents to perceive each other through their mutual interactions and their similarity, as estimated by the Euclidean distance between their neural network parameters. Data visualized for three- and four-neuron circuits. All agents are good at recognizing themselves, as can be seen by the points in the upper left corner. Some pairs of agents are good at recognizing each other and others are not, as can be seen by the spread of points in the upper right quadrangle. (**B**) Histogram of the performance of all three- and four-neuron pair of circuits. The distribution is slightly bimodal, suggesting that most pairs are either good at finding each other or fail to find each other.

of agents were relatively good at finding each other, while other pairs were not good at all. Among the 3-neuron circuits, performance ranged from 0.478 to 1 with a mean of 0.779 ± 0.202 (Figure 7B). Among the 4-neuron circuits, performance ranged from 0.511 to 1 with a mean of 0.841 ± 0.194 (Figure 7D). Importantly, the distribution of performance across the pair of agents was bimodal for both circuit sizes (Figure 8B). While both 3-neuron and 4-neuron circuits could maintain mutual interaction with non-clonal agents, the 4-neuron circuits were moderately better, more consistently achieving perfect performance.

We note that the performance matrix for both circuit sizes is not fully symmetric, suggesting that an agent's position on different sides of the 1-dimensional ring environment can influence their performance. This is especially clear with, for example, 3-neuron circuits #s 8 and 9. When circuit #8 is on one side of the ring, the pair achieves performance $= 1$. When circuit #9 is on the other side of the ring, the pair's performance worsens to $= 0.673$. Determining what exactly drives the asymmetry of this interaction is beyond the scope of this paper.

Does the similarity of a pair of agents' parameters for the neural controller predict how well they will perform together? To address this, we computed the normalized Euclidean distance between every pair of 3-neuron agents and every pair of 4-neuron agents (Figure 7A and C). Our first step was to visualize the relationship between the performance at finding each other between all of these pairs of agents and their similarity, as given by the Euclidean distance (Figure 8A). From visual inspection, there appears to be little or no relationship between the two. There are examples of pairs of agents that are equally close to each other or equally far apart in terms of their parametric distances, and yet they differ dramatically in terms of their ability to detect each other. A statistical analysis of the pairs revealed there is a relationship between proximity in parameter space and performance for both 3-neuron circuits ($r = -0.515, p < 0.001$) and 4-neuron circuits ($r = -0.266, p = 0.011$; Figure 7A). Because there are many kinds of distance metrics, we also computed the cosine similarity of all pairs of agents' parameters. Cosine similarity yielded similar results ($r = 0.53, p < 0.001$ for 3-neurons; $r = 0.271, p = 0.009$ for 4-neurons). Given that both samples were relatively small ($N = 9$ and $N = 13$ for 3-neuron and 4-neuron circuits respectively), we present these findings with caution, especially since the relationship weakens with the larger 4-neuron sample.

A similarity matrix can be cast as a network, which opens the door to analytical tools from network science. To better understand how parameter similarity relates to performance, we clustered the cosine similarity matrices for 3-neuron and 4-neuron circuits. For each circuit size, we ran 1000 iterations of the Louvain algorithm of modularity maximization (Newman & Girvan, 2004). Each iteration yielded a partition of the matrix into communities. Using the community assignments from, each partition, we computed the coassignment probability or the proportion of partitions in which two agents were assigned to the same community on the basis of the similarity of their parameters. Neither circuit size yielded a significant relationship between coassignment probability and performance ($r = 0.301, p = 0.044$ for 3-neuron circuits and $r = 0.123, p = 0.284$ for 4-neuron circuits). This result suggests that agents do not consistently perform well with their neighbors in parameter space. Furthermore, it casts doubt on the underpowered correlations between performance and euclidean distance and cosine similarity. It is possible that the asymmetry of performance muddies the correlation.

Given that clonal pairs presented different dynamical motifs, is there a relationship between dynamical features and performance between non-clonal pairs? We approached this question from many angles, but we focus on the presence of limit cycles in clonal pairings. First, we identified whether the non-clonal agents were the same in that they both had or both did not have a limit cycle in their clonal dynamics, but this did not predict performance in non-clonal pairs ($\chi^2(76) = 81.971, p = 0.3$ for 3-neuron and $\chi^2(121) = 120.16, p = 0.5$ for 4-neuron circuits). Second, we tested whether the presence of a limit cycle in agent 1's clonal dynamics predicted non-clonal performance, but it did not ($\chi^2(76) = 78.571, p = 0.397$ for 3-neuron and $\chi^2(121) = 108.53, p = 0.785$ for 4-neuron circuits). Because the performance matrix is asymmetric, we also tested whether a limit cycle in agent 2's clonal dynamics predicted non-clonal performance, but it did not ($\chi^2(76) = 76.562, p = 0.46$ for 3-neuron and $\chi^2(121) = 113.1, p = .683$ for 4-neuron circuits). All in all, there was no relationship between clonal dynamics and non-clonal performance. We note that we use clonal dynamics to predict non-clonal performance, instead of using non-clonal dynamics to predict non-clonal performance, because the non-clonal dynamics involve twice as many dimensions. This increase in dimensionality makes the analysis much more difficult, and therefore outside of the scope of this paper.

We have noted the different strategies that evolved agents use to solve the perceptual crossing task, but is there a relationship between perpetual versus minimal crossing and performance? We structure these tests in the same way as our tests of dynamical motifs and performance above, examining whether (1) the sameness of strategy, (2) the first agent's strategy, or (3) the second agent's strategy predicts performance. None of these tests for either circuit size yield significant results (all $\chi^2(76)s < 83.751$ for 3-neurons, all $\chi^2(121)s < 122.4$ for 4-neurons, and for both all ps > 0.2539 for both). Clonal solutions, then, do not predict non-clonal performance.

Although it is interesting to find a relationship between inter-agent performance and more abstract spaces like parameter similarity or features in the non-autonomous dynamics of the circuits, it is entirely possible that the only factor that is informative about whether an agent can detect another agent is purely behavioral. This might be particularly true if the agents are not truly good at detecting mutual interactions, but are more merely detecting certain, for example, frequencies and amplitudes of movement. Accordingly, our final question in this analysis concerns this issue: Are these agents detecting a mutual interaction or merely a certain frequency and amplitude of movement? To address this, we performed a psychophysical experiment in which we introduced a 'decoy' object that moves with a set frequency and amplitude. We used the same agent for which we presented in-depth dynamical analyses (a 3-neuron perpetual crosser with a limit cycle). We varied both the amplitude and frequency from 0 to 1.0. Because frequency and amplitude were fixed within a given trial, the decoy had no way of sensing the agent and no way of changing its behavior to respond to the agent. Despite this, the agent is ‘tricked’ by the decoy for a large number of frequency-amplitude settings. Interestingly, the number of crossings the agent and decoy achieve is highly discontinuous across the space, and the agent is not especially successful when the decoy’s frequency and amplitude are set to match its own (Figure 9). While there is a range of frequencies and amplitudes where the agent is successful, this range does not include its own approximate frequency and amplitude.

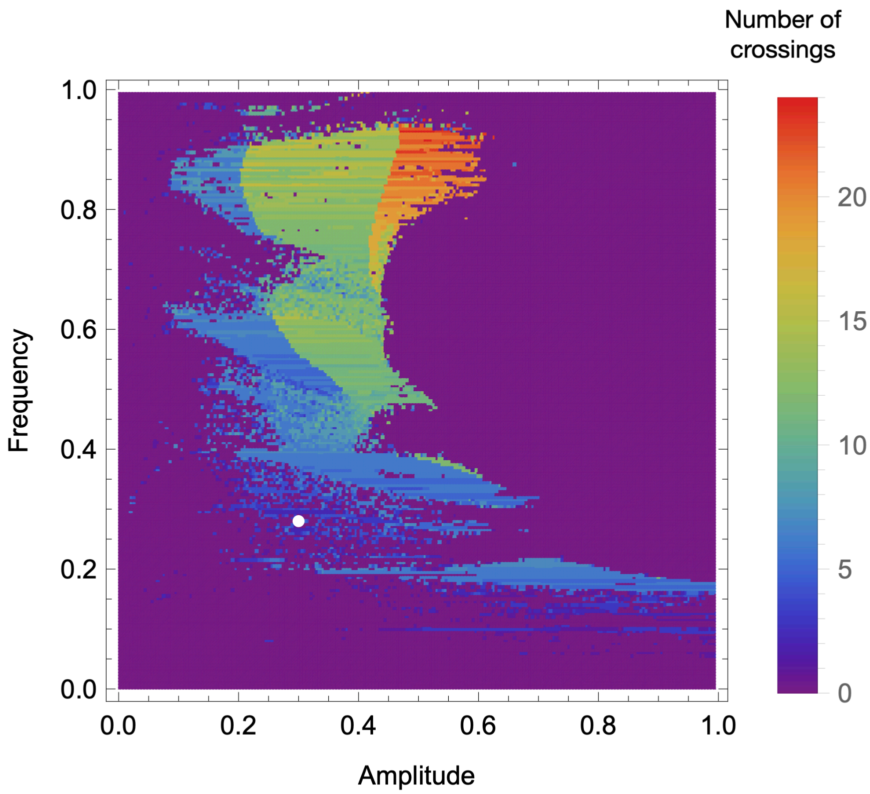

Figure 9: Psychophysical experiment. We further examined the solution analyzed in detail in the previous sections in an environment without other agents, shadows, or fixed objects, but instead with a ‘decoy’: an agent-like object that moved left and right at a certain frequency and a certain amplitude. For any one trial, the frequency and amplitude of the decoy were fixed over time. Thus, the decoy had no way of sensing the agent and no way of changing its behavior in relation to it. We counted the number of times the agent being examined crossed the decoy for a range of different frequencies and amplitudes. The white disk represents the estimated typical amplitude and frequency of an interaction based on the recorded behavior of an agent interacting with its clone, cf. Figure 4(iii).

4 Discussion

We extended the perceptual crossing simulations to have no sensory delay, to have a continuous sensor, to disallow cheating by relying on fixed objects or shadows, to test the robustness of successful agents, and to test successful agents against other successful agents. As in our previous work (Izquierdo et al., 2022), we observed both perpetual and minimal crossing strategies among successful robust solutions. Our extensive robustness testing provides more rigorous solutions to the perceptual crossing task.

Through dynamical analyses of some of the perpetual solutions, we clarified the neural dynamics underlying interaction. Crucially, limit cycles present in individual dynamics do not seem to play a key role. Instead, what drives successful behavior is the switching between the different directions of movement as the sensor is activated. While we have investigated the neural dynamics of simpler evolved model agents (3 neurons), it is important to acknowledge that we are only observing a 3-dimensional slice of the much larger 7-dimensional space (4 dimensions if the sensor is included). While we have concluded that limit cycles at the individual level do not seem to play a key role in perceptual crossing, the existence of larger dimensional limit sets in state space, at this point, can not be ruled out.

In our inter-agent analysis, we observed that many of the successful robust solutions were relatively good at detecting other agents. Whether any of the properties of any individual agent, such as parameter similarity, predict high performance is inconclusive. Correlations between parameter similarity and performance were technically significant, but the effect sizes and p-values are less impressive in the larger sample of 4-neuron agents. Moreover, there was no relationship between coassignment probability and performance. Disentangling the drivers of successful non-clonal performance may involve digging deeper into the dynamics and psychophysics of a non-clonal pair. Regardless, this analysis was the first to show that the perceptual crossing task does not require clonal agents (but see Froese and

Di Paolo, 2009 for a disruption of the symmetry of neural controllers). While using clones is an effective simplifying assumption to make analyses more tractable, that it is not necessary highlights the opportunities for dialogue between virtual and human instances of perceptual crossing.

Through psychophysical analysis, we showed the limitations of the solutions and therefore of the task setup. Although we analyzed only one agent, we found that it achieved high-quality performance for only a small range of frequencies and amplitudes. Interestingly, this range did not include its own frequency and amplitude. Whether this result holds for other successful robust agents is an open question.

4.1 Future Work

Although our work has attempted to clarify the nature of many of the constitutive components that comprise perceptual crossing simulations, there are still many open questions and avenues for additional work.

A key feature that has helped to define perceptual crossing is the valuable dialogue that has been established between empirical experiments and simulation studies in artificial life. There are many areas that we find can and should, be implemented in perceptual crossing experiments. In the experimental literature, researchers have explored the effect of ”previously recorded behavior” from another person to use as a ”decoy” in perceptual crossing experiments (Iizuka et al., 2012; Lenay & Stewart, 2012). While our psychophysical results have demonstrated a limitation in the task setup for artificial agents, the effect of decoys of all sorts (eg. random walkers or oscillators at a certain frequency) could be explored in a more systematic fashion. While there has been some attention to the nature of the sensory delay in perceptual crossing experiments in addition to our previous work, a more deliberate investigation of the sensory delays behavior effects in both simulation and experiment is needed (Iizuka et al., 2015; Izquierdo et al., 2022). Additionally, while we have explored the larger behavioral repertoire of evolved agents (perpetual vs minimal crossers) a similar clarification has not been seen in human subjects. Further work in perceptual crossing experiments may look towards what conditions could be studied in human tasks for systematically investigating different behavioral strategies similar to what we've discovered here. What conditions in the task lead to one behavior achieving more success than another?

The replacement of the discrete sensor with a continuous sigmoidal sensor has opened up many opportunities for dynamical systems analysis of perceptual crossing simulations. Further work can explore the role that larger dimensional structures play in coupled agents. For example, the presence of toroidal dynamics in coupled oscillators has been well characterized in dynamical systems (Verhulst, 2015). Whether or not something similar exists between coupled evolved agents is an open question. While we have shown that the presence of limit cycles in individual neural dynamics does not predict the success of an agent, we have not been able to conclude the same with higher dimensional structures. Further work in dynamical systems analysis of perceptual crossers can explore the role that larger dimensional structures play in coupled agents. The existence and characterization of these structures are as of yet unknown. Additionally, it is important to acknowledge that we have only studied the dynamics of *perpetual* crossers, but the dynamics of minimal crossers also are important to interrogate in a rigorous manner.

Finally, we tested already evolved circuits with non-clonal pairs. It is unclear whether it is possible to evolve successful perceptual crossers that only experiences circuits different from their own. Such a setup more closely mirrors human social interaction but it might approach the boundary of what artificial agents are capable of.

References

Auvray, M., Lenay, C., & Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas in Psychology*, *27*(1), 32–47.

Beer, R. D. (1995). On the dynamics of small continuous-time recurrent neural networks. *Adaptive Behavior*, *3*(4), 469–509.

Beer, R. D. (1996). Toward the evolution of dynamical neural networks for minimally cognitive behavior. *From Animals to Animats*, *4*, 421–429.

Cikara, M., Martinez, J. E., & Lewis, N. A. (2022). Moving beyond social categories by incorporating context in social psychological theory. *Nature Reviews Psychology*, *1*(9), 537–549.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10), 441–447.

Di Paolo, E. A. (2000). Behavioral coordination, structural congruence and entrainment in a simulation of acoustically coupled agents. *Adaptive Behavior*, *8*(1), 27–48.

Di Paolo, E. A., Rohde, M., & Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? modelling the dynamics of perceptual crossing. *New Ideas in Psychology*, *26*(2), 278–294.

Froese, T. (2018). Searching for the conditions of genuine intersubjectivity: From Agent-Based Models to Perceptual Crossing Experiments. In *The Oxford Handbook of 4E Cognition*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198735410.013.9

Froese, T., & Di Paolo, E. A. (2008). Stability of coordination requires mutuality of interaction in a model of embodied agents. *International Conference on Simulation of Adaptive Behavior*, 52–61.

Froese, T., & Di Paolo, E. A. (2009). Toward minimally social behavior: Social psychology meets evolutionary robotics. *European Conference on Artificial Life*, 426–433.

Froese, T., & Di Paolo, E. A. (2010). Modelling social interaction as perceptual crossing: An investigation into the dynamics of the interaction process. *Connection Science*, *22*(1), 43–68.

Froese, T., & Di Paolo, E. A. (2011). The enactive approach: Theoretical sketches from cell to society. *Pragmatics & Cognition*, *19*(1), 1–36.

Froese, T., & Fuchs, T. (2012). The extended body: A case study in the neurophenomenology of social interaction. *Phenomenology and the Cognitive Sciences*, *11*(2), 205–235. https://doi.org/10.1007/s11097-012-9254-2

Iizuka, H., & Di Paolo, E. A. (2007). Minimal agency detection of embodied agents. *European Conference on Artificial Life*, 485–494.

Iizuka, H., & Ikegami, T. (2004). Adaptability and diversity in simulated turn-taking behavior. *Artificial Life*, *10*(4), 361–378.

Iizuka, H., Marocco, D., Ando, H., & Maeda, T. (2012). Turn-taking supports humanlikeness and communication in perceptual crossing experiments — toward developing human-like communicable interface devices, 1–4. https://doi.org/10.1109/VR.2012.6180953

Iizuka, H., Saitoh, S., Marocco, D., & Yamamoto, M. (2015). Time delay effect on social interaction dynamics. *Proceedings of the 3rd International Conference on Human-Agent Interaction*, 217–219.

Ikegami, T., & Iizuka, H. (2007). Turn-taking interaction as a cooperative and co-creative process. *Infant Behavior and Development*, *30*(2), 278–288.

Izquierdo, E. J., Severino, G. J., & Merritt, H. (2022). Perpetual crossers without sensory delay: Revisiting the perceptual crossing simulation studies. *ALIFE 2022: The 2022 Conference on Artificial Life*.

Lenay, C., & Stewart, J. (2012). Minimalist approach to perceptual interactions. *Frontiers in Human Neuroscience*, *6*, 98.

Lewis Jr, N. A. (2022). What would make cognitive science more useful? *Trends in Cognitive Sciences*.

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, *69*(2), 026113.

Pfeiffer, U. J., Timmermans, B., Vogeley, K., Frith, C. D., & Schilbach, L. (2013). Towards a neuroscience of social interaction.

Quinn, M. (2001). Evolving communication without dedicated communication channels. *European Conference on Artificial Life*, 357–366.

Reséndiz-Benhumea, G. M., & Froese, T. (2020). Enhanced neural complexity is achieved by mutually coordinated embodied social interaction: A statespace analysis. *Second International Workshop on Agent Based Modelling of Human Behaviour*.

Reséndiz-Benhumea, G. M., Sangati, E., Sangati, F., Keshmiri, S., & Froese, T. (2021). Shrunken social brains? a minimal model of the role of social interaction in neural complexity. *Frontiers in Neurorobotics*, *15*, 72.

Rohde, M., & Paolo, E. D. (2008). Embodiment and perceptual crossing in 2d. *International Conference on Simulation of Adaptive Behavior*, 83–92.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience 1. *Behavioral and Brain Sciences*, *36*(4), 393–414.

Verhulst, F. (2015). Torus break-down and bifurcations in coupled oscillators. https://doi.org/10.13140/RG.2.1.2394.1602

Williams, P. L., Beer, R. D., & Gasser, M. (2008). Evolving referential communication in embodied dynamical agents. *Eleventh International Conference on the Simulation and Synthesis of Living Systems*, 702–709.

A Supplementary Materials

A.1 Agent and Neural Controller

The behavior of each agent is controlled by a continuous-time recurrent neural network (Beer, 1995) with the following state equation:

$$\tau_i \dot{y}_i = -y_i + \sum_{j=1}^{N} w_{ji}\sigma(y_j + \theta_j) + g_i s + I_i \tag{1}$$

where y_i is the state of each neuron, τ is the time constant, w_{ji} is the strength of the connection from the j^{th} to the i^{th} neuron, θ is a bias term, $\sigma(x) = 1/(1 + e^{-x})$ is the standard logistic activation function, g_i is the sensory weight from the sensor s to neuron i, and I_i represents an external input to each neuron. The output of a neuron is $o_i = \sigma(y_j + \theta_i)$.

In the simulation, the objects (i.e., the agent's avatars, the agents' shadows, and the static objects) occupy approximately 2 units of space. An agent's sensor is activated when an object is close enough to it, according to the following equation (see Fig. 1B):

$$s(d) = \frac{1}{1 + e^{8(d-1)}} \tag{2}$$

where d is the absolute distance between the midpoint of the agent and the midpoint of the other object.

Following the original simulation studies (Fig. 1A), the sensor, s, is fully connected to all neurons in the circuit via a set of weights; the neurons are fully interconnected (including self-connections); and two of the neurons are chosen to drive the left and right motors, respectively. The velocity of an agent is proportional to the difference between the outputs of the two motor neurons: $v = \gamma(o_1 - o_2)$, where o_1 and o_2 represent the outputs of the neuron controlling the left and right motors, respectively, and γ is a constant that determines the agent's maximum possible velocity. The maximum velocity was set to $\gamma = 2$.

A.2 Evolution and Fitness Function

The neural parameters of the controller are evolved using a real-valued genetic algorithm. Given that during evolution both agents are clones of each other in terms of their neural controller, each genome encodes the parameters for only one neural controller. The following neural parameters, with corresponding ranges, are evolved: time-constants $\tau \in [1, 10]$, biases $\theta \in [-8, 8]$, and all connection weights (from sensors to neurons, g, and between neurons, w) $\in [-8, 8]$. We used a generational algorithm with rank-based selection and a population size of 96 genotypes. Successive generations are formed by first applying random Gaussian mutations to each parent genome with a mutation variance of 0.05 (see Beer, 1996 for details). In addition, uniform crossover is applied with 50% probability. A child replaces its parent if its performance is greater than or equal to that of the parent; otherwise the parent is retained.

The goal of the fitness evaluation is to get agents to find each other. Since the avatars, shadows, and fixed objects are indistinguishable to either agent, success in this task requires that the agents evolve a system for accurately detecting *mutual* interactions. We evaluate the performance of a pair of agents by systematically varying the starting location of the two agents. Specifically, the starting location for the first agent in a pair is chosen between 0 and 600 in steps of 50; the starting location for the second agent in the pair is between 0 and the first agent's starting location, for a total of 78 trials. Each trial lasts 800 time units and proceeds as follows. First, the neural states of both agents are initialized to 0. During the first 400 units of time, the agents interact without evaluation. We treat this as a transient period because it allows for agents initialized at the maximum starting distance moving at their maximum velocity enough time to traverse the ring environment and find each other. Finally, for the remainder of the simulation after the transient period, we record and normalize the distance between the two agents. For a given trial, the score that a given pair of agents with a given neural controller receive is:

$$f = 1 - \frac{\bar{d} - 2}{298} \tag{3}$$

where $\bar{d}$ is the average separation between the two agents during a trial (excluding the initial transient period), 298 is the maximum spatial distance between the two agents. Since the 1-D environment wraps around between 0 and 600 units, 300 is the maximum spatial distance between points on the ring; and because the agents are 2 units wide and the sensors are binary in the previous studies, the agents cannot detect proximity beyond 2 units of space away from each other. The final fitness of the evaluation is the average fitness across all trials. Note that the fitness is normalized to run between 0 and 1 based on the minimum distance at which an agent can sense the other agent. Also, the fitness evaluation is deterministic: the starting positions of the agents are deterministic, the position of the fixed objects does not change, and the relative position of both shadows to the agent is fixed.