Data Preprocessing

The raw fish behavior logs were subjected to a series of preprocessing steps for consistency across logs. First behavior terms were normalized and classified in which like terms and behaviors were grouped, for example the "pot entry" and "pot exit" were normalized into the single term "pot". Behaviors such as pecking at and chasing other fish were classified as aggressive behaviors, fleeing as aversive, while circling, following and pot were classified as reproductive behaviors . The percentage of time each fish spent in a given behavior for a trial was calculated relative to the total observation duration.

Markov Chain Graph Construction

The core analysis involved constructing Markov chain graphs to visualize the transitions between different behavioral states. The Markov chain is a probabilistic model that allows the representation of behavioral transitions as a directed graph. Each node in the graph represented an individual behavioral state, while the size of each node corresponded to the percentage of time a fish spent in that specific behavior. The directed edges between nodes illustrated the transitions from one behavior to another. The thickness of each edge was determined by the percentage of transitions that occurred between the two connected behaviors relative to all transitions with a percentage threshold of 0.001, highlighting the relative importance of each transition.

To construct the Markov chain graphs, a transition probability matrix was computed from the preprocessed data. Each entry (i, j) in the matrix represented the probability of transitioning from behavior i to behavior j. The matrix was then used to generate the graph representation of the behavioral transitions using graph visualization libraries or software.