The causal role of synergy in collective problem-solving

Ketika Garg[1*†], Cody Moser[2,3*†], Hannah Dromiack[4,5], Zara Anwarzai[6], Gabriel Ramos-Fernandez[7,8*]

[1]Department of Humanities and Social Sciences, California Institute of Technology, Pasadena, 91125, CA, USA.

[2]Department of Cognitive and Information Sciences, University of California, Merced, 95343, CA, USA.

[3]School of Collective Intelligence, Université Mohammed VI Polytechnique, Rabat, 11103, Morocco.

[4]BEYOND Center for Fundamental Concepts in Science, Arizona State University, Tempe, 85281, AZ, USA.

[5]Department of Physics, Arizona State University, Tempe, 85281, AZ, USA.

[6]Department of Philosophy, Simon Fraser University, Vancouver, V6B 5K3, British Columbia, Canada.

[7]Research Institute on Applied Mathematics and Systems, National Autonomous University of Mexico, Mexico City, 04510, Mexico.

[8]Global Research Centre for Diverse Intelligences, University of St Andrews, St Andrews, KY16 9AJ, United Kingdom.

*Corresponding author(s). E-mail(s): kgarg@caltech.edu; cmoserj@gmail.com; gabriel@aries.iimas.unam.mx;

Contributing authors: hdromiac@asu.edu; zanwarza@sfu.ca;

†These authors contributed equally to this work.

Keywords: Network structure, Redundancy, Synergy, Collective problem-solving, collective intelligence

1 Significance Statement

Collective intelligence, or the ability of groups to solve problems more effectively than individuals, is a widely-occurring phenomenon in the natural world. Research has shown that a group's structure influences its ability to solve complex problems. However, the causal mechanisms underlying why some groups perform better than others remain unclear. Utilizing information-theoretic measures, we demonstrate that synergy of information processing—where individuals combine knowledge in novel ways—can explain the relationship between network structure and group performance. We find that while network structures generally predict performance, all of them can perform equally well when their information processing is synergistic. Our findings suggest synergy, rather than structure alone, is the fundamental driver of success in collectively intelligent systems.

2 Abstract

The influence of network structure on a system's capacity to solve complex problems is a central focus in collective intelligence research. However, the causal mechanisms through which structure shapes system-level outcomes remain poorly understood. Prior work has examined the relationship between network structure and performance at a coarse-grained level, paying limited attention to how solutions depend on information dynamics within groups. To explore these dynamics, we utilize an agent-based model, the Potions Task, which operationalizes problem-solving as a combinatorial process and is thereby ideal for studying collective problem-solving through information-theoretic analyses. We examine information-based metrics at the level of agent pairs across the network, analyzing their variation over time to determine how they predict problem-solving across network structures. Specifically, we measure redundancy (or conversely, synergy) of solutions discovered by agents with respect to the network's global knowledge at a given time. While we replicate a well-established finding that small-world networks support efficient problem-solving, our results reveal a deeper mechanism where small-world networks achieve efficiency by balancing local redundancy with long-range synergy. Furthermore, we find synergy to be a consistent predictor of group performance, including in networks typically considered inefficient. Synergy in information processing, measured at both local and global levels, therefore mediates the effects of structure and can override them entirely, implicating it as a fundamental determinant of group performance. By introducing a causal framework for information processing in collectives, our study delivers a more definitive explanation of collective problem-solving than those offered by structural analysis alone.

3 Introduction

The discipline of collective intelligence has brought to light different ways that individuals work together to achieve group-level success. Phenomena including swarm intelligence, distributed cognition, collective decision-making, adaptive systemic behavior, emergent coordination, among others, have been studied under the umbrella of collective intelligence, showing commonalities in group-level behavior across diverse morphologies and interaction strategies [1–5]. A wide range of studies have inquired about the effects of network structures [6, 7], agent strategies [8, 9], and agent diversity [10, 11] as driving factors underlying collective problem-solving. Inquiries have been targeted at a range of phenomena, including how individuals learn from other individuals [12, 13], how they complement each other [14], and how groups as a whole possess information not possessed by any individual member of the group [15, 16]. However, despite an increased understanding of the role that diversity and structure play in collective problem-solving, causal-mechanistic explanations [17, 18] of why specific structures lead to more optimal outcomes remain underexplored.

Collective problem-solving, a specific form of collective intelligence, refers to a process wherein individual components in a collective interact to solve a given problem, such as finding new nest sites [19] or foraging for food [20]. Such tasks require managing the trade-off between exploration for new solutions and exploitation of existing solutions. One way for groups or networks to achieve this is to solve different parts of the problem in a distributed manner, similar to parallel problem-solving [6, 21–25]. In a seminal study, Lazer et al. [6] showed that network structure plays a key role in facilitating (or preventing) this type of advantageous distributed problem-solving. The speed, or efficiency, of information diffusion across a network affects how that information is processed to generate new solutions. In complex problem-solving tasks, there exists a trade-off between finding similar solutions to those which are known and exploration of the problem space to find novel ones[6]. Network features such as the degree of clustering or the average path length can contribute to how information is shared, and thus, the efficiency with which a network finds a solution [6]. Empirical studies [26–29] have further supported similar effects of the structure of interactions on a group's capacity for problem-solving.

This effect of certain system-level arrangements on performance at the group-level has been called to attention through aggregate descriptions of a system's state. Such approaches tend to focus on individual performance metrics (e.g., efficiency). These studies have only sparsely examined causal mechanisms underlying performance, such as how information is distributed and integrated in a network to solve a given problem [11]. The focus on final outcomes has made it difficult to understand the underlying causal mechanisms of collective problem-solving—that is, which processes within a collective behavior give rise to particular outcomes [17, 18]. Despite the fact that the literature on collective-problem solving employs causal-mechanistic language—for example, describing how divided populations sub-divide work [30] or how core-periphery arrangements allow for the *exploration* versus *exploitation* of given solutions [31]—causal explorations of these general dynamics remain sparse.

We address this gap by using the Potions Task model [32, 33] that helps simulate complex and collective problem-solving in networked agents, while applying information to analyze the underlying mechanisms of information processing [14, 34]. Methods from information theory have been and are often applied to quantify communication and causal interactions across disciplines (e.g. ecology [35], animal behavior [36], neuroscience [37, 38], genetics [39, 40] and collective action [41]). They provide a model-independent and probabilistic approach to studying how information is encoded, shared, and integrated between different parts of a system.

In the Potions Task, agents in a network share and combine information with their neighbors to discover new "potions". The problem in the task is solved when agents successfully create a *crossover potion*, which requires integrating discoveries from two distinct solution trajectories. These trajectories represent the different sequences of intermediate potions needed to reach a final solution. Depending on the network structure and the order of discoveries, agents may follow different pathways to reach a solution, either sequentially finding potions along each trajectory ($A \to B$ or $B \to A$; shown as Pathways 2 and 3 in Fig. 1) or by balancing the explore-exploit trade-off and advancing along both trajectories in parallel (A and B; shown as Pathway 1 in Fig. 1). Previous work in this task found that partially connected groups outperform fully connected groups because full connectivity accelerates convergence to suboptimal solutions—a pattern also observed by [6] in an NK landscape problem-solving. Similarly, Migliano et al. [33], using a real-world implementation of the Potions task, showed that structured clustering helps balance problem-solving speed and exploration of multiple solution trajectories. Because the potions task requires agents to combine partial solutions over time according to a structured rule set, it is particularly well-suited to investigate collective information processing through information theory. Its compositional structure, where agent performance and inventories encode their history of interactions, enables a unique decomposition of information processing at both the individual and group levels. It also approximates key features of canonical NK landscape models, which are commonly used to study collective performance on complex tasks in networks [6, 11, 42].

Our information-theoretic framework (similar to partial information decomposition [43]) enables us to differentiate between the distinct ways in which a system's components contribute to a collective's search for solutions (Fig. 2). The information that agents have about the global knowledge can be *redundant* (when agents share overlapping knowledge) or *synergistic* (when they hold complementary knowledge that enables new discoveries). In addition, *complementarity* captures how agents collectively cover the problem space. Low levels of redundancy (conversely, high levels of synergy) and high complementarity can make systems *collectively intelligent* in the sense that no individual component can solve the problem on its own, but can do so collectively [14, 41]. Applied to our context, the compositional nature of the Potions Task enables the use of such metrics to analyze how knowledge is transmitted, distributed, and recombined over time across the network. Each time an agent discovers a new potion, it is added to their inventory and shared with neighbors, generating a

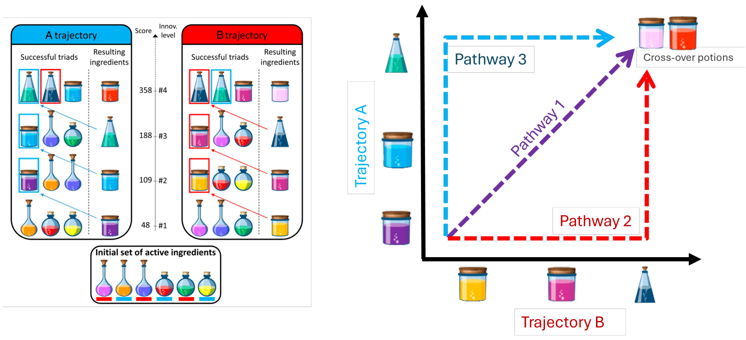


Fig. 1: *Left:* In the Potions Task, each agent starts with an initial set of three active ingredients (out of the initial set: bottom). As the simulation progresses, they combine existing potions with their neighbors to generate new ones. This innovation can occur along two trajectories (A and B), with potions from both trajectories being combined, eventually reaching either of two crossover potions (top). The figure is replicated from Derex & Boyd [32]. *Right:* Possible solution pathways that a network can traverse to discover new potions and eventually find the cross-over potions. They can either simultaneously solve along A and B (Pathway 1) or they can sequentially solve along A first and then B (Pathway 3) or along B first and then A (Pathway 2).

structured record of how knowledge flows and accumulates in the system. To examine how network properties shape this information processing and problem-solving, we simulate different configurations of agent interactions, using the small-world formalism [44] to create networks with different clustering coefficient and average path length. This approach allows us to systematically investigate how the structure of interactions influences the distribution and integration of knowledge across the collective.

Our study aims to elucidate the causal mechanisms of collective problem-solving by pursuing three primary objectives: (1) quantifying the relationship between network structure and collective performance, (2) investigating how network topology shapes the distribution of information among agents, and (3) determining how redundancy (conversely, synergy) and complementarity of information processing affects group-level problem solving. We predict that network structures mediate agent-level redundancy, and that the less redundant that information is, the better a given network will be at collective problem solving. Ultimately, our work bridges the gap between static network metrics and the dynamics of information-processing and helps to uncover the causal link between network structure and collective problem-solving.

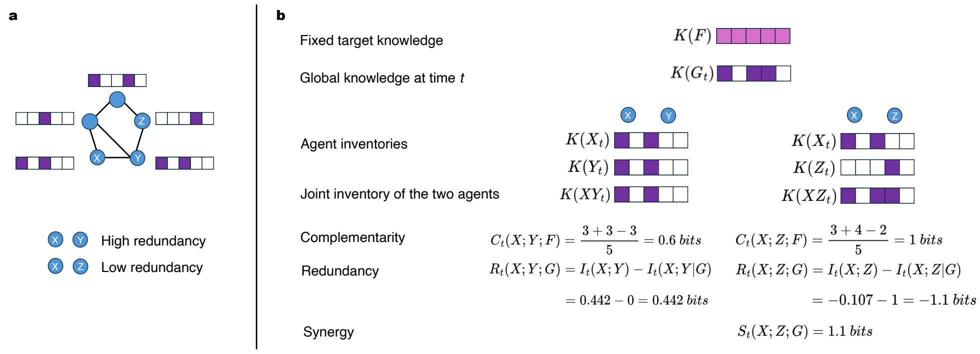


Fig. 2: Illustration of agent inventories in the model and the information-theoretic measures applied to them. a) Representation of a network of agents and their inventories. Each blue circle represents an agent (only X, Y and Z labeled), with their inventories shown as 5-bit binary arrays (for graphical purposes, simplified from the actual 14-bit model). Purple squares indicate potions known to each agent and white squares potions that are unknown so far. We can tell that X and Y are more redundant in their knowledge than X and Z based on the overlap between their respective inventories. b) Calculations of the information-theoretic measures. Two reference vectors are the full target set of potions, $K(F)$, or ”complete knowledge”, represented by pink squares; and the “global knowledge” at time t held by the collective, $K(G_t)$. Comparing the same two pairs of agents as in a), we define three information-theoretic measures: complementarity (C), redundancy (R) and synergy (S). C measures how much closer agents get to $K(F)$ after combining their inventories. The pair (X, Z) is more complementary than (X, Y), as their combined inventory $K(XZ)$ covers more of $K(F)$ with minimal overlap of known potions. R quantifies shared knowledge between agents; (X, Y) exhibit high redundancy since they contribute same potions to the global knowledge, $K(G_t)$. If R is negative, we consider S to be the absolute value of R. $S > 0$ captures cases when the pooled information from a pair predicts the network's information $K(G_t)$ in a way that neither agent alone can (as in the case of X and Z). These measures are calculated at each time step across all agent pairs to track knowledge distribution in the network over time.

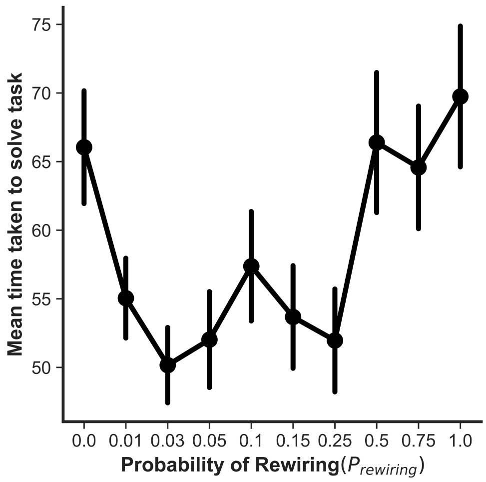


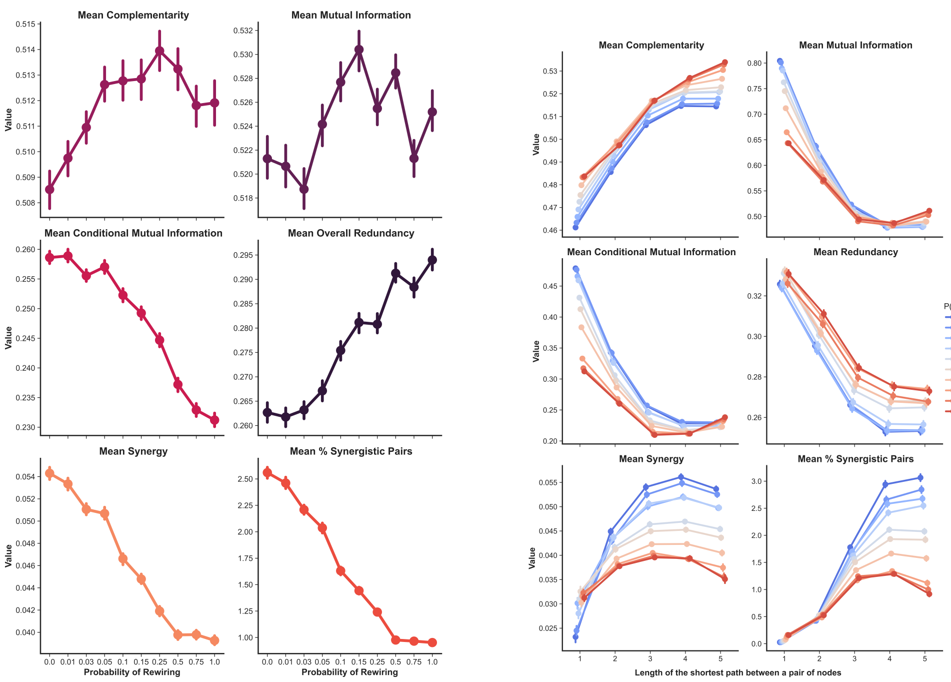
Fig. 3: Performance of networks as a function of the rewiring parameter $P_{rewiring}$. Longer time taken to finish the task indicates lower performance. Error bars correspond to the standard error.

4 Results

4.1 Small-world networks are more efficient at complex problem-solving

We measure the performance of a network based on the time steps it takes to reach either of the two crossover potions, which are made by combining the terminal potions from both trajectories (Fig.3). We used the small-world formalism to construct networks with different structures [44]. The algorithm begins with a regular network, where nodes in a ring are connected to each of their four neighboring nodes, and then changes to the small-world and random regimes by rewiring existing edges at random based on a given probability of rewiring ($P_{rewiring}$, see Fig. S1).

We found that networks with low rewiring probabilities ($P_{rewiring} = 0$ - 0.01) exhibit slower and more variable task completion times, due to their lattice-like structure with high local clustering and long path lengths. At intermediate probabilities of edge rewiring (0.03 - 0.25), the network structure transitions towards the small-world regime, balancing local clustering with shorter path lengths. This configuration leads to improved performance reflected in faster task completion times. Performance worsens as the networks transition from the small-world regime towards more random configurations (0.5-1.0). These results suggest that small-world properties may promote efficient problem-solving in the Potions Task.



(a) (b)

Fig. 4: Average information-theoretic metrics at the pairwise level (a) across all pairs of nodes as a function of $P_{rewiring}$ and (b) as a function of path length for different values of $P_{rewiring}$. Values are means and bars show standard errors across each simulation. Complementary measures how much two agents cover the full possible set of potions. Mutual Information, Conditional Mutual Information, Redundancy and Synergy measure how much information a pair of agents contain about the potions discovered by the whole network. % of Synergistic Pairs refers to the percentage of agent pairs with synergy (i.e., negative redundancy values). Each metric was aggregated at each step of the simulation and for each pair of agents in a network.

4.2 Network structure determines the distribution and extent of redundant or synergistic information

For each iteration and network type, we computed the pairwise values of complementarity, mutual information, conditional mutual information, redundancy and synergy, as well as the proportion of synergistic pairs. Each metric was calculated at each time step across all agent pairs to track knowledge distribution in the network. We aggregated these measures across each time step and between each pair of nodes to capture global patterns (Fig. 4a).

We found that complementarity (C) towards the complete set of potions in the task increases as networks transition from regular lattices (low $P_{rewiring}$) to the small-world regime, and decline again in the random regime (Fig. 4a, top row). This trend indicates that nodes in small-world networks hold more diverse and non-overlapping knowledge about the complete target set. In contrast, agents' inventories become more redundant or fragmented in regular or random-like networks. C is computed using Hamming distance, which measures how far an agent's inventory is from the complete set of potions. We assess how much unique knowledge they contribute by comparing the Hamming distances of individual agents with their combined inventories. High C means that a pair of nodes expands their individual discovered knowledge in combination, and low C implies that their information is either redundant or incomplete.

Mutual information ($I(X;Y)$) likewise peaks in the small-world regime, suggesting that information in one node is less predictable from another in small-world structures. However, informational overlap increases as networks become too ordered or random. Conditional mutual information ($I(X;Y;G)$, Fig. 4a, middle-left) quantifies the dependence of a pair of nodes' knowledge on the network's overall knowledge. We found that this measure decreases steadily with increasing randomness, indicating that inventories are more interdependent in structured networks and more independent in random networks.

Redundancy (R) increases as networks become more random (Fig. 4a, middle row), while both mean synergy and the proportion of synergistic pairs decline (Fig. 4a, bottom row). Synergy (S) captures a special case of low redundancy, where a pair of nodes together provide more information about the network's global knowledge that neither can provide alone. This effect suggests that small-world networks balance unique contributions and shared knowledge, which may be important for more efficient problem-solving where coordination is necessary.

To understand how network structure contributes to the amount of redundancy observed, we assessed the distribution of information throughout the networks by comparing the average (rather than pairwise) values of the information measures between pairs at different path lengths (Fig. 4b). Our results show that nodes closer to each other (i.e., with shorter path lengths) are more redundant (and, correspondingly, less complementary and synergistic) than nodes farther away. The degree of redundancy in far away nodes depends on the $P_{rewiring}$: distant nodes in regular-like networks (shown in blue) are much more synergistic than distant nodes in random-like networks (shown in red).

Taken together, these results demonstrate that redundancy in information processing is affected by a network's structure. Next, we assess how redundancy in information processing changes over time within a network and whether it can explain network performance.

4.3 Redundant information processing leads to inefficient problem-solving

We have shown that structure influences how information is processed and distributed by agents across networks. However, aggregated measures can mask the transient dynamics and transitions that could reveal how different structures process information over time. To observe how these metrics change over time, we calculated them between each pair of nodes at each time step. To meaningfully compare the dynamics between different networks with different run times, we normalized each time step in an iteration by the total number of time steps the iteration took to find the solution. Thus, the normalized time goes from 0, the initial step, to 1, the time step it took each network to find the solution or maximum time step allowed in the simulation (approximately 0.6% of iterations did not find a solution within 1000 time steps).

The results in the previous section suggest that network structure, as defined by the rewiring parameter, is an important determinant of performance. However, the potions task model contains inherent levels of stochasticity in potion creation. For example, due to the random nature of potion combination at the earliest step, separate sub-graphs can begin climbing separate, or the same, trajectories, leading to either redundant or synergistic outcomes. To investigate whether information processing differs between efficient and inefficient networks, we divided the networks into performance categories based on the time taken to find the solution. Specifically, we created three bins of equal size based on the total steps required to solve the task: fast (2-334 steps), medium (338-664 steps) and slow (667-999 steps).

Our results show that complementarity decreases over time across all categories and networks (Fig. 5 (A) top-row). This trend can be explained by the fact that as agents get increasingly closer to the complete target set, there are fewer opportunities for complementary pairings of knowledge. In fast networks, complementarity steadily decreases towards task completion, but the overall magnitude of the decrease is lower than in the medium and slow performing networks (top right two plots). The fast networks sustained higher values of C throughout the task indicating that agents do not find the same potions immediately and therefore are less redundant, maintaining their diversity for longer. In contrast, medium and slow performing networks exhibit a steeper or more varied decline in C.

We also observed that the dynamics in medium and slow networks differ based on the network structure. Regular-like networks (low $P_{rewiring}$), with high clustering and long path lengths, show the steepest decline in C early in the simulation. This implies that communities of regular networks find new potions quickly but fail to combine them to generate crossover potions due to a lack of connections between communities. On the other hand, random-like networks exhibit a more gradual decline in C because their global connectivity enables agents to transfer (and mix) information more effectively (i.e. faster and with a wider information distribution) and generate new potions. Small-world networks lie in the intermediate range, striking a balance between the local clustering of regular networks and the global connectivity of random networks.

We observe a similar overall trend in redundancy (R) values over time (Fig. 5 (A), middle row). At the start of the simulation, all agents have completely redundant inventories. Over time, in fast networks, R decreases as agents discover different

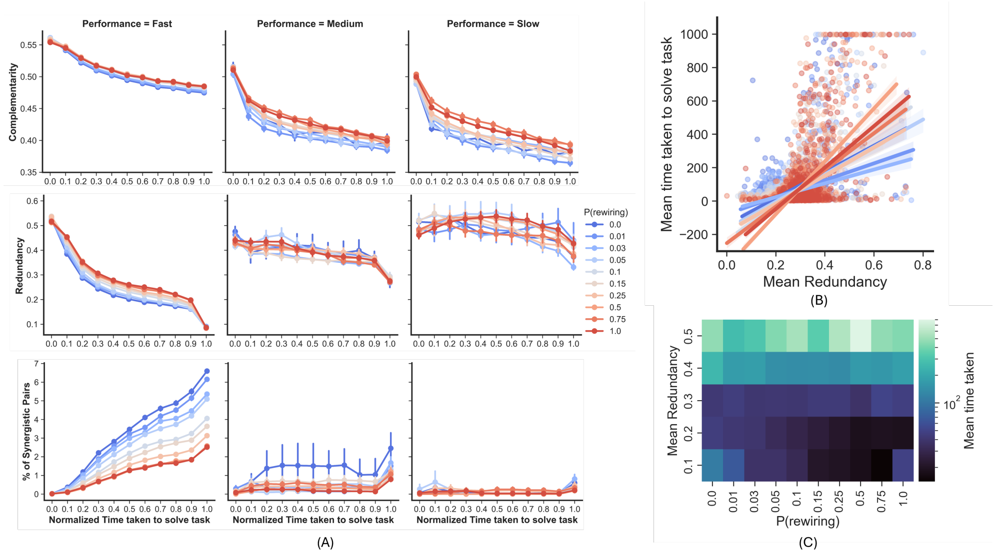


Fig. 5: (A) Dynamics of complementarity, redundancy and % of synergistic pairs in fast (*left*), medium (*center*), and slow (*right*) networks. Each metric is averaged across all pairs of nodes for each network at each normalized timestep. Error bars indicate standard error. (B) Correlation between mean redundancy and time taken to solve the task for a given $P_{rewiring}$ (shown in different colors). Each point corresponds to mean redundancy value of each iteration. (C) Mean time taken to solve the task (indicated by color) across each $P_{rewiring}$ and mean redundancy (binned). The color map is normalized on a logarithmic scale to account for the skewed data distribution. Lighter colors indicate longer simulation times and inefficient problem-solving.

subsets of potions, which implies they begin to contribute different potions to the network's current global knowledge. In contrast, in medium and slow networks, R values remain relatively stable or even increase, as in random networks. This effect suggests that in inefficient networks, agents tend to discover the same subsets of potions simultaneously. As a result, their inventories remain highly similar over time.

Conversely, the average percentage of synergistic pairs increases over time, particularly in the fast-performing networks, whereas medium and slow networks have very few or no synergistic pairs throughout the simulation (Fig. 5 (A), bottom row). We also observe that fast regular-like networks accumulate the highest number of synergistic pairs. This high level of synergy in regular networks implies that different communities in the graph contribute different subsets of potions to the global knowledge. On the other hand, fast random networks accrue the lowest number of synergistic pairs over time, due to low clustering and a high rate of information exchange.

To further investigate the effect of redundancy on network performance, we correlated the mean redundancy values with the time taken to complete the task (Fig. 5 (B)). The scatter plot shows a positive correlation between redundancy and the time taken, which further demonstrates that redundancy decreases network efficiency. We also created a heatmap to assess whether this effect holds true across different network structures (Fig. 5 (C)). We find that for low redundancy values (lower rows), all networks tend to finish the task efficiently. Networks closer to the random regime (0.5-1.0) appear to perform especially well at low redundancies compared to networks closer to the regular regime (0.0-0.01). Factors leading to low redundancy may be especially advantageous in random networks, given their structural predisposition to be highly redundant (as shown in Fig. 4a).

These results show that redundancy in information processing can help explain network performance, where high levels of redundancy leads to inefficient networks. In the next section, we investigate whether these patterns at the node pair level also emerge at the network level.

4.4 Efficient networks solve the task synergistically

Thus far, our results suggest that synergy in information processing at the agent pair level is essential for efficient problem-solving. If agents in a network generate the same potions (i.e., redundant information), then the network is slow to finish the task. However, information theory metrics at the level of agent pairs fail to show us how synergistic information processing leads to solution (i.e., potions) generation at the network-level.

In the potions task, the task can primarily be solved through two pathways, as shown in Fig. 1. We conceptualized the problem-space as a 4 x 4 two-dimensional grid, where one dimension represents potions along trajectory A and the other represents trajectory B. To visualize the pathways that networks take to solve the task, we represented potion generation along a trajectory as movement in the problem-space. We computed stepwise change in potions generated over the two trajectories as movement vectors (a, b). We then aggregated the vectors per grid cell to derive the flow or path of the movement and its intensity ($\sqrt{\Sigma(a^2 + b^2)}$), as shown in Fig. 6(A).

The vector field shows that in the fastest networks (Fig. 6 (A; left most)), the solutions tend to follow a diagonal path, meaning the system finds potions along both trajectories simultaneously (i.e. Pathway 1 in Fig. 1). In networks performing at a medium pace (Fig. 6 (A; middle)), the concentration of "discovery" along this diagonal path decreases, while the likelihood of alternate paths—those that initially follow A (or B) before switching to B (or A)—increases. Finally, in the slowest networks (Fig. 6 (A; right)), the diagonal path is favored the least, and the trajectories appear to split off or branch toward either A or B and simultaneous information processing seems unlikely. These effects can be seen in Fig. 7, as well. Nodes in the fast network (top-row) solve for A (blue color) and B (red color) potions simultaneously. In the example of a medium network, the proportion of B potions appears to be lower than A potions, and in the slow network, B potions appear to dominate the network.

We also quantified the diagonality of the pathways (see Fig. S2) by calculating the angular deviation of a path from an ideal diagonal and normalizing them such that perfect diagonal movement along the problem space would be scored 1 and deviations

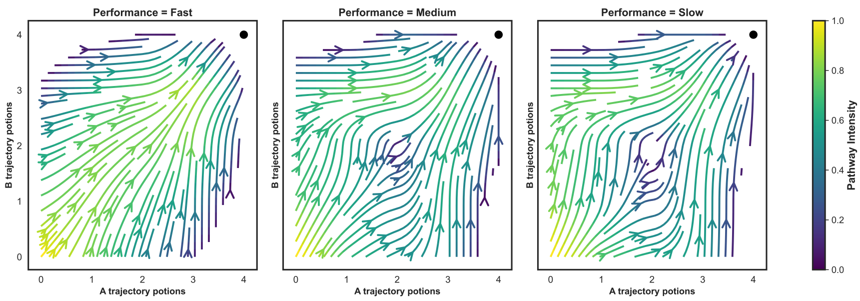


Fig. 6: The 4×4 grid represents the problem-space of the task, where potions must be found along the two trajectories (depicted along the two axes). This streamline plot visualizes the discovery patterns of networks as they progress toward the crossover potion (black dot, upper right). Discovery over the problem-space is derived from stepwise changes in potion generation along each trajectory, aggregated to represent directional flow. Color intensity (logarithmic scale) reflects pathway frequency, with a threshold applied on the intensity to mask low-intensity paths. The plot reveals that fast networks are more likely to advance along both trajectories simultaneously (high intensity along the diagonal), whereas medium and slow networks exhibit less balanced progression (lower diagonal intensity).

from it would be less than 1. We found that the paths in fast networks that solve the problem are almost perfectly diagonal, irrespective of the $P_{rewiring}$.

These results are consistent with our previous findings which show that synergy leads to efficient problem-solving in the task. Synergy at the level of pairwise connections allows the networks to compute different parts of the problem-space and achieve crossover potions quickly, which translates to parallel problem-solving (i.e., Pathway 1 in Fig. 1). On the other hand, high levels of redundant interactions lead to networks finding new potions along one trajectory first and then along the other. Such serial processing (i.e., Pathway 2 and 3 in Fig. 1) renders networks inefficient. Thus, by analyzing information processing at both the pairwise and network levels, our results offer a comprehensive picture of how synergy plays a causal role in network performance.

5 Discussion

Network science has revolutionized the study of collective behavior by offering a powerful framework for analyzing interactions within groups. A large body of theoretical and empirical work has demonstrated that network structure plays a crucial role in shaping how effectively groups solve problems (rev. in [7, 45]). In this paper, we extend this literature by investigating information processing as a causal mechanism [17, 18] that links network structure to collective performance. We show that group problem-solving efficiency is mediated by information processing dynamics at both the pairwise and network levels. In the Potions Task, where agents must combine discoveries from

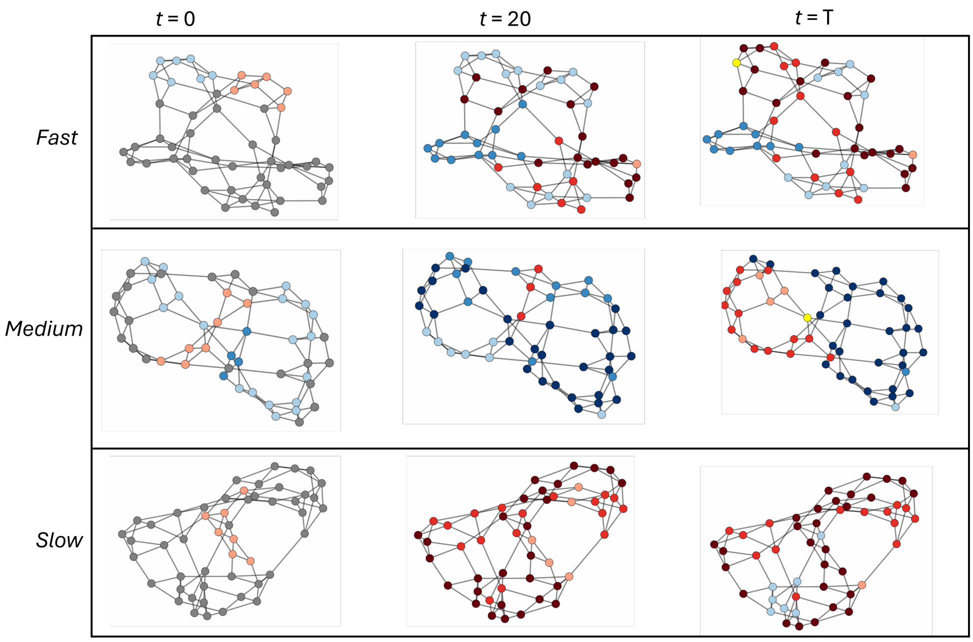


Fig. 7: Dynamics of potion generation in an example of each fast (top row), medium (middle row) and slow (bottom row) networks. The two potion trajectories (A and B) in the model are depicted using blue and red. To reach the crossover potions, three potions must be discovered along each trajectory. The darkness or intensity of a color depicts the n^{th} potion found along a given trajectory by a node (e.g., light blue corresponds to **1a**, medium blue to **2a**, dark blue to **3a**). The color yellow shows the crossover potions. All networks have $P_{rewiring} = 0.1$. Three time steps for each network are shown: $t = 0$ (when they begin the task); $t = 20$ (after 20 time steps); $t = T$ (the final step). The fast network shows a distribution of both red and blue hues, and the slow network is dominated by red hues, suggesting greater redundancy than in fast networks. Animated versions of this figure can be found in the Supplementary materials.

two trajectories to create crossover potions, low redundancy emerges as a central determinant of success. Redundant agent pairs share overlapping inventories already common in the network, limiting exploration and reducing the chance of discovering critical crossover potions. In contrast, synergistic pairs contribute unique, complementary potions that enhance collective search. At the network level, high redundancy leads to slower, serial problem-solving, and low redundancy enables parallel exploration. These results highlight low redundancy (and conversely high synergy) as core causal mechanisms linking network structure to collective performance.

Previous studies (e.g., [6, 44, 46–48]) have shown that small-world networks often outperform regular (high clustering, long paths) and random (low clustering, short paths) networks in a variety of tasks. Our results identify redundancy in information processing as a key causal mechanism underlying this advantage. We find that in regular networks, limited long-range connections inhibit information exchange, leading to low redundancy and fragmented exploration. Random networks, by contrast, facilitate rapid information spread, but their low clustering often results in high redundancy, causing premature convergence on sub-optimal solutions. Small-world networks, which combine high clustering with short path lengths, balance local specialization through redundancy with an agent's nearest neighbors and global integration through complementarity across longer path lengths.

In addition to identifying redundancy as a mechanism linking network structure to system-level outcomes, we also show how low redundancy itself can override the structural constraints imposed by networks. This effect implicates synergy not merely as an intermediary, but as a fundamental determinant of group performance, even in cases where structural network measures fail to account for observed outcomes. Although small-world networks are generally efficient, other network types *can* perform well when they process information synergistically rather than redundantly, an outcome enabled by noise or the stochastic nature of the Potions Task. For instance, random-like networks, despite their structural predisposition toward redundancy, can achieve high performance when agents contribute complementary information, allowing fast and synergistic exchange along short paths. As Cantor et al. [49] note, network topology alone does not determine collective outcomes, as transmission dynamics and stochasticity also shape a system's "realized connectivity." Our findings reinforce this view, showing that while network structure influences performance by moderating redundancy, low redundancy itself provides a more precise explanatory framework than structure alone.

Our results also build on and extend general research highlighting the role diversity plays in collective problem-solving [10, 11, 50, 51]. By contributing a formal, information-theoretic framework linking diversity to redundancy and quantifying how information processing dynamics shape collective performance, we show that low redundancy serves as a unifying mechanism behind many important observed effects of network structure, helping to formalize prior notions of "transient diversity" [11, 52, 53]. Causally, we find that the network structures that promote diversity are those that provide access to a wider range of potential solutions (i.e., low redundancy in potions found), formally quantified through our measures of redundancy and synergy in agent interactions. Furthermore, rather than simply benefiting from initial diversity, high-performing networks sustain complementary contributions over time and exhibit a shallower decline in synergy than slower networks (Fig. 5A). This dynamic suggests that performance depends not only on network structure, but on the persistence of diversity throughout the problem-solving process.

Recently, researchers have begun to re-evaluate foundational questions in collective behavior by evaluating the causal mechanisms underlying information processing in collective systems. This trend reflects a growing effort to move beyond phenomenological correlations and towards more causal explanations of how network structure, agent behavior, and transmission dynamics interact to shape collective outcomes. For example, Guilbeault & Centola [54] showed that standard network metrics, such as path length and centrality, fail to capture key aspects of complex information diffusion in networks. Similarly, Barkoczi & Galesic [24] demonstrated that network outcomes are influenced not only by topology, but mediated by agent-level strategies for integrating information, such as different social learning heuristics. This growing body of literature underscores the need for metrics that take into account how information is causally processed within networks. Our results highlight a link between agent-level information processing, structural dynamics, and group-level performance and represent one step towards bridging causal mechanisms across scales in collective behavior.

Future research on collective intelligence should move towards a more comprehensive understanding of how information is processed across multiple levels of organization, ideally examining redundancy in relation to trade-offs with other forms of information processing and connecting insights from formal models to real-world systems. This direction would include exploring the benefits of low redundancy in maintaining system robustness and coordination, and in preventing information loss. While our findings emphasize how redundancy can limit exploration, prior work has shown that it also promotes resilience by providing common ground and buffering against disruption [55–59].

Furthermore, findings from the modeling literature should be increasingly aligned with insights drawn from real-world contexts. Context-dependent information processing [60] and group features such as diversification, complementarity and skill dependency have been identified as critical drivers of collective performance [61–64]. A mature science of collective intelligence will require greater attention to how task structure interacts with agent-level strategies, such as learning [24], memory [65], and communication [66, 67], that shape the nature of information processing itself. Investigating these strategies may offer a critical pathway for bridging the gap between abstract models and the complex dynamics observed in real-world systems.

Collective intelligence is a field with interest across diverse disciplines and systems, ranging from bacterial colonies and fish schools to human teams. Such diverse perspectives come with a diversity in measures to quantify collective intelligence, which can lead to fragmented approaches towards the phenomenon, hampering comparative inquiries. Our work advances this effort by applying an information-theoretic approach to quantify a form of collective intelligence [14, 38, 41], offering a step towards uncovering mechanistic fundamentals of collective intelligence.

6 Methods

6.1 Model Description

1. **Model initialization:** A network is generated according to the Watts–Strogatz model using a specified rewiring probability (Fig. S1). For each node of the graph, an agent is initialized with a score of zero and an inventory comprised of six potions: three from an **A trajectory** and three from **B trajectory**. Each inventory is

comprised of three parts: the name/level of the potion (e.g., a1, a2, a3, b1, b2, b3) and a score which each initial potion and potions discovered thereafter carries for itself (with scores of 6, 8, and 10 for the three initial potions or ingredients in each trajectory).

2. **Dyad selection:** At each step, each agent chooses a partner they are connected to on the network with a random probability. As neighbors are simply chosen with a random probability, it is possible for a focal neighbor to select an individual which is already interacting with them (e.g., if a network is initialized with just two agents, the two agents will simply select each other).
3. **Potion selection:** In the model, new potions are formed by triad combinations of old potions. As triad combinations are made between dyads of agents, the focal agent randomly selects whether it will be trading either one or two potions with their partner. The focal agent and its partner then cycle through their respective inventories, assigning probabilities to each potion in the array. This probability is obtained by summing the innovation scores of each potion and dividing individual scores by each sum (e.g., the initial inventory innovation scores of 6, 8, 10, 6, 8, 10 will yield respective probabilities of .125, .167, .208, .125, .167, .208).
4. **Potion combination:** Agents and their partners then select the number of potions previously assigned to them in the last step, based on potions' calculated probabilities and without replacement, and combine them. The combination is saved as a list and compared to lists of valid combinations copied directly from Derex and Boyd [32] (SI Appendix, section 1). If an invalid combination is made, nothing happens. If a valid combination is made, then the dyad adds a new innovation (with its own respective score) to their inventories.
5. **Innovation diffusion:** If a new potion is added to the agents' inventories, both agents then check the inventories of all of their partners and spread it to neighbors which do not already possess it.
6. **Scoring:** Scores are then obtained for each agent based on the tier of discovery an agent has obtained: with the first tier yielding a score of 48, second tier 109, third tier 188, and the fourth tier (which requires a crossover from the A and B trajectory) being 358. The maximum score of a potion in an agent's list is determined to be their overall score.
7. **End and Crossover:** The simulation ends either when it has reached 1000 steps or when the network has achieved a *crossover event*, whereby the final inventions in both trajectories are themselves combined, indicating the network has discovered and united both paths of exploration.

6.2 Network Configurations

We systematically varied the configurations in which agents interact using synthetic networks built according to a simple and well-known process [45]. All networks contain 50 agents. We took advantage of the simplicity in [44]'s model for exploring the full range of network configurations, from regular to random networks. We start with a regular network, where nodes are positioned in a ring and are connected only to their four closest neighbors. Then, with probability $P_{rewiring}$, we take a link and rewire it, connecting one of the original nodes with a randomly chosen node anywhere in the network. This is repeated for all links in the network. Watts et al. showed that, for initially regular networks (relatively low values of $P_{rewiring}$) contains enough long-distance connections to maintain the initially high clustering coefficient, C, but has a significantly lower path length, L (both of which define the small-world property) [44]. For higher values of $P_{rewiring}$, C decreases to the point of being indistinguishable from a random network. We used the Watts-Strogatz graph algorithm in the *NetworkX* package in Python, with a degree of 4 (see Fig. S1). Rewiring probabilities were set to 0, 0.01, 0.03, 0.05, 0.1, 0.15, 0.25, 0.50, 0.75 and 1.0.

6.3 Information theoretic tools

We represent each agent's inventory as a 14-bit binary array, indicating whether each potion is present or absent.

6.3.1 Redundancy and Synergy via Mutual Information

To quantify the level of redundant information in a network, we compute pairwise mutual information between all pairs of agents at each time step. We then compare how redundant their information is with respect to the global knowledge of the network at a given time step, t. We define the global knowledge (or collective information) of the network as all the unique potions that have been discovered by any agent in the network by time t.

Mutual information $I_t(X;Y)$ tells us the amount of information that can be gained about the inventory of agent X by knowing the information possessed by agent Y. Information known to an agent is defined through Shannon Entropy $H_t(X)$, which captures how much an agent knows about the full potion set, where $p_t(x)$ is the probability associated with each potion x of the full target set at time t.

$$\begin{aligned} I_t(X;Y) &= H_t(X) + H_t(Y) - H_t(X,Y) \\ &= \sum_{x\in X}\sum_{y\in Y} p_t(x,y)\log\frac{p_t(x,y)}{p_t(x)p_t(y)} \\ \text{where } H_t(X) &= -\sum_{x\in X} p_t(x)\log p_t(x) \end{aligned} \tag{1}$$

To quantify redundancy, R, we use the following:

$$\begin{aligned} R_t(X;Y;G) &= I_t(X;Y) - I_t(X;Y|G), \\ \text{where } G_t &= \bigvee_{i=1}^{n} x_{i,t} \end{aligned} \tag{2}$$

G_t represents the global knowledge of the network at time t, and $I_t(X;Y|G)$ is the conditional mutual information between X, Y, and G at time t. Conditional mutual information quantifies the conditional dependence of two variables on a third one. Depending on the level of dependence, it can be higher or lower than the mutual information between X and Y, and it can, thus, yield a negative or positive value of R.

Following partial information decomposition [43], R represents the overlap or redundancy in the information contained in X and Y about the global knowledge, G. Negative values of R imply a synergistic relationship, where X and Y *together* provide information about G that can only be known when X and Y are considered simultaneously.

To assess the level of redundancy in the network, we compute the average value of R_t over time and across all agent pairs. Additionally, we define a synergy measure S, which captures the extent of synergistic relationships by considering only the negative values of R_t:

$$S_t = \frac{1}{N} \sum_{\substack{X,Y \\ R_t(X;Y;G)<0}} |R_t(X;Y;G)| \tag{3}$$

where N is the total number of agent pairs with negative R_t. We also compute the frequency of synergistic pairs, which represents the percentage of pairs in the network exhibiting synergy at each time step:

$$P_t = \frac{|\{(X,Y) \mid R_t(X;Y;G) < 0\}|}{|\{(X,Y)\}|} \tag{4}$$

where $|\{(X,Y)\}|$ denotes the total number of agent pairs in the network. A higher S_t or P_t indicates stronger synergistic interactions in the network. In the context of the potions task, if a pair of agents knows different potions that have been discovered by the collective network, it would imply that they contribute to the global knowledge in a synergistic manner.

6.3.2 Complementarity via Hamming Distance

While the mutual information approach quantifies how agents contribute to the global knowledge at each time step, we also assess how complementary two agents' knowledge is in relation to the full target set. This measure constitutes an alternative, simple way to measure how efficiently agents cover a finite and small set of potions.

Each agent's inventory can be represented as a binary vector of size 14, where each position is 1 if the agent possesses the corresponding potion and 0 otherwise. The full target set F is represented as a vector of 1's of size 14, indicating that all potions are present. The Hamming distance between an agent's inventory X and the full set F measures how much information the agent lacks:

$$D_t(X,F) = \sum_{i=1}^{z=14} |x_{i,t} - f_i| \tag{5}$$

where $x_{i,t} - f_i \in \{0,1\}$ for all $i \in \{1,\ldots,14\}$, and t denotes the time step. A higher value means that the agent is missing more potions or more information about the target set.

To measure how complementary two agents' knowledge is at each time step, we compare their individual Hamming distances to the full set with the Hamming distance of their combined knowledge:

$$C_t(X,Y) = \frac{D_t(X,F) + D_t(Y,F) - D_t(X \vee Y, F)}{14} \tag{6}$$

where $D_t(X,F)$ and $D_t(Y,F)$ are the distances of agents X and Y from the full target set at time t, and $D_t(X \vee Y, F)$ is the distance of their combined knowledge (bit-wise OR operation). This computation is performed at each time step and for every pair of agents in the network. A higher $C_t(X,Y)$ indicates that agents contribute non-overlapping knowledge, meaning they are more complementary in their knowledge about the full set. Over time, this measure allows us to track how complementarity evolves in the network as agents acquire new information.

7 Acknowledgments

We thank the Diverse Intelligences Summer Institute for bringing together this group of researchers. This work received support from the Institute for Humane Studies Grant No. IHS017741 (to CM) and CONACYT CF-2019-263958 and UNAM-PASPA (to GRF). We also thank Thomas Varley and Jesse Milzman for feedback on information-theoretic measures used in this study, and Luke Leckie, Yoonjung Yi, and members of the Smaldino lab at UC Merced for their thoughts during discussions on this project.

Competing interests. Authors declare no competing interests.

Author Contributions. C.M. designed the model and generated data. K.G. analyzed data and generated figures. K.G, G.R.F., and C.M. wrote the paper. K.G., H.D., and G.R.F. conceptualized and implemented the information-theoretic measures. All authors contributed to the conceptualization of the study and the editing of the paper.

Code Availability. The agent-based model, information theory calculations, and code to replicate figures are available at a GitHub repository (link).

References

[1] Wolpert, D. H. & Tumer, K. A survey of collective intelligence (1999).

[2] Krause, J., Ruxton, G. D. & Krause, S. Swarm intelligence in animals and humans. *Trends in Ecology & Evolution* **25**, 28–34 (2010).

[3] Trigo, P. & Coelho, H. *Collective-Intelligence and Decision-Making*, 61–76 (Springer Netherlands, Dordrecht, 2011).

[4] Malone, T. W. & Bernstein, M. *Handbook of collective intelligence* (MIT press, 2022).

[5] Garnier, S. & Moussaïd, M. We the swarm—methodological, theoretical, and societal (r) evolutions in collective decision-making research. *Collective Intelligence* **1**, 26339137221133400 (2022).

[6] Lazer, D. & Friedman, A. The network structure of exploration and exploitation. *Administrative Science Quarterly* **52**, 667–694 (2007).

[7] Centola, D. The network science of collective intelligence. *Trends in Cognitive Sciences* **26**, 923–941 (2022).

[8] Zafeiris, A. & Vicsek, T. Group performance is maximized by hierarchical competence distribution. *Nature Communications* **4**, 2484 (2013).

[9] Mann, R. P. & Helbing, D. Optimal incentives for collective intelligence. *Proceedings of the National Academy of Sciences* **114**, 5077–5082 (2017).

[10] Hong, L. & Page, S. E. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* **101**, 16385–16389 (2004).

[11] Smaldino, P. E., Moser, C., Pérez Velilla, A. & Werling, M. Maintaining transient diversity is a general principle for improving collective problem solving. *Perspectives on Psychological Science* **19**, 454–464 (2024).

[12] Kameda, T., Toyokawa, W. & Tindale, R. S. Information aggregation and collective intelligence beyond the wisdom of crowds. *Nature Reviews Psychology* **1**, 345–357 (2022).

[13] Collet, J. *et al.* Mechanisms of collective learning: how can animal groups improve collective performance when repeating a task? *Philosophical Transactions of the Royal Society B* **378**, 20220060 (2023).

[14] Bettencourt, L. M. The rules of information aggregation and emergence of collective intelligent behavior. *Topics in Cognitive Science* **1**, 598–620 (2009).

[15] Hinsz, V. B., Tindale, R. S. & Vollrath, D. A. The emerging conceptualization of groups as information processors. *Psychological Bulletin* **121**, 43 (1997).

[16] Moussaid, M., Garnier, S., Theraulaz, G. & Helbing, D. Collective information processing and pattern formation in swarms, flocks, and crowds. *Topics in Cognitive Science* **1**, 469–497 (2009).

[17] Craver, C. F. *Mechanistic Explanation* (MIT Press, 2025). Https://oecs.mit.edu/pub/vgigt1aq.

[18] Braillard, P.-A., Malaterre, C., Breidenmoser, T. & Wolkenhauer, O. Explanation and organizing principles in systems biology. *Explanation in Biology: An Enquiry into the Diversity of Explanatory Patterns in the Life Sciences* 249–264 (2015).

[19] Seeley, T. D. Consensus building during nest-site selection in honey bee swarms: the expiration of dissent. *Behavioral Ecology and Sociobiology* **53**, 417–424 (2003).

[20] King, A. J. & Sueur, C. Where next? group coordination and collective decision making by primates. *International Journal of Primatology* **32**, 1245–1267 (2011).

[21] Levinthal, D. A. Adaptation on rugged landscapes. *Management Science* **43**, 934–950 (1997).

[22] Aral, S. & Van Alstyne, M. The diversity-bandwidth trade-off. *American Journal of Sociology* **117**, 90–171 (2011).

[23] Garg, K., Kello, C. T. & Smaldino, P. E. Individual exploration and selective social learning: balancing exploration–exploitation trade-offs in collective foraging. *Journal of the Royal Society Interface* **19**, 20210915 (2022).

[24] Barkoczi, D. & Galesic, M. Social learning strategies modify the effect of network structure on group performance. *Nature Communications* **7**, 13109 (2016).

[25] Sloman, S. J., Goldstone, R. L. & Gonzalez, C. A social interpolation model of group problem-solving. *Cognitive Science* **45**, e13066 (2021).

[26] Daniels, B. C. & Romanczuk, P. Quantifying the impact of network structure on speed and accuracy in collective decision-making. *Theory in Biosciences* **140**, 379–390 (2021).

[27] Almaatouq, A. *et al.* Adaptive social networks promote the wisdom of crowds. *Proceedings of the National Academy of Sciences* **117**, 11379–11386 (2020).

[28] Centola, D. *et al.* Experimental evidence for structured information–sharing networks reducing medical errors. *Proceedings of the National Academy of Sciences* **120**, e2108290120 (2023).

[29] Baumann, F., Czaplicka, A. & Rahwan, I. Network structure shapes the impact of diversity in collective learning. *Scientific Reports* **14**, 2491 (2024).

[30] Moser, C. & Smaldino, P. E. Innovation-facilitating networks create inequality. *Proceedings of the Royal Society B* **290**, 20232281 (2023).

[31] Milzman, J. & Moser, C. Decentralized core-periphery structure in social networks accelerates cultural innovation in agent-based model. *arXiv preprint arXiv:2302.12121* (2023).

[32] Derex, M. & Boyd, R. Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences* **113**, 2982–2987 (2016).

[33] Migliano, A. B. *et al.* Hunter-gatherer multilevel sociality accelerates cumulative cultural evolution. *Science Advances* **6**, eaax5913 (2020).

[34] Shannon, C. E. A mathematical theory of communication. *The Bell System Technical Journal* **27**, 379–423 (1948).

[35] MacArthur, R. Fluctuations of animal populations and a measure of community stability. *Ecology* **36**, 533–536 (1955).

[36] Ramos-Fernández, G. *et al.* Quantifying uncertainty due to fission–fusion dynamics as a component of social complexity. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20180532 (2018).

[37] Luppi, A. I., Rosas, F. E., Mediano, P. A., Menon, D. K. & Stamatakis, E. A. Information decomposition and the informational architecture of the brain. *Trends in Cognitive Sciences* (2024).

[38] Varley, T. F., Pope, M., Grazia, M., Joshua & Sporns, O. Partial entropy decomposition reveals higher-order information structures in human brain activity. *Proceedings of the National Academy of Sciences* **120**, e2300888120 (2023).

[39] Schneider, T. D. Information content of individual genetic sequences. *Journal of Theoretical Biology* **189**, 427–441 (1997).

[40] Chan, T. E., Stumpf, M. P. & Babtie, A. C. Gene regulatory network inference from single-cell data using multivariate information measures. *Cell Systems* **5**, 251–267 (2017).

[41] Kemp, J. T., Kline, A. G. & Bettencourt, L. M. Information synergy maximizes the growth rate of heterogeneous groups. *PNAS Nexus* **3**, pgae072 (2024).

[42] Kauffman, S. & Levin, S. Towards a general theory of adaptive walks on rugged landscapes. *Journal of theoretical Biology* **128**, 11–45 (1987).

[43] Williams, P. L. & Beer, R. D. Nonnegative decomposition of multivariate information. *arXiv preprint arXiv:1004.2515* (2010).

[44] Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world'networks. *nature* **393**, 440–442 (1998).

[45] Newman, M. E. The structure and function of complex networks. *SIAM review* **45**, 167–256 (2003).

[46] Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* **10**, 186–198 (2009).

[47] Uzzi, B. & Spiro, J. Collaboration and creativity: The small world problem. *American Journal of Sociology* **111**, 447–504 (2005).

[48] Fontanari, J. F. & Rodrigues, F. A. Influence of network topology on cooperative problem-solving systems. *Theory in Biosciences* **135**, 101–110 (2016).

[49] Cantor, M. *et al.* Social network architecture and the tempo of cumulative cultural evolution. *Proceedings of the Royal Society B* **288**, 20203107 (2021).

[50] Stephen, A. T., Zubcsek, P. P. & Goldenberg, J. Lower connectivity is better: The effects of network structure on redundancy of ideas and customer innovativeness in interdependent ideation tasks. *Journal of Marketing Research* **53**, 263–279 (2016).

[51] Cao, Y. *et al.* Effects of network connectivity and functional diversity distribution on human collective ideation. *npj Complexity* **2**, 2 (2025).

[52] Zollman, K. J. The epistemic benefit of transient diversity. *Erkenntnis* **72**, 17–35 (2010).

[53] Wu, J. & O'Connor, C. How should we promote transient diversity in science? *Synthese* **201**, 37 (2023).

[54] Guilbeault, D. & Centola, D. Topological measures for identifying and predicting the spread of complex contagions. *Nature Communications* **12**, 4430 (2021).

[55] Walker, M. A. Redundancy in collaborative dialogue. *arXiv preprint cmp-lg/9503017* (1995).

[56] Artime, O. *et al.* Robustness and resilience of complex networks. *Nature Reviews Physics* **6**, 114–131 (2024).

[57] Tononi, G., Sporns, O. & Edelman, G. M. Measures of degeneracy and redundancy in biological networks. *Proceedings of the National Academy of Sciences* **96**, 3257–3262 (1999).

[58] Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *Nature* **406**, 378–382 (2000).

[59] Hillary, F. G. & Grafman, J. H. Injured brains and adaptive networks: the benefits and costs of hyperconnectivity. *Trends in Cognitive Sciences* **21**, 385–401 (2017).

[60] Steiner, I. D. *Group Processes and Productivity* (Academic Press, New York, 1972).

[61] Neffke, F. M. The value of complementary co-workers. *Science Advances* **5**, eaax3370 (2019).

[62] Neffke, F. & Henning, M. Skill relatedness and firm diversification. *Strategic Management Journal* **34**, 297–316 (2013).

[63] Hosseinioun, M., Neffke, F., Zhang, L. & Youn, H. Skill dependencies uncover nested human capital. *Nature Human Behaviour* 1–15 (2025).

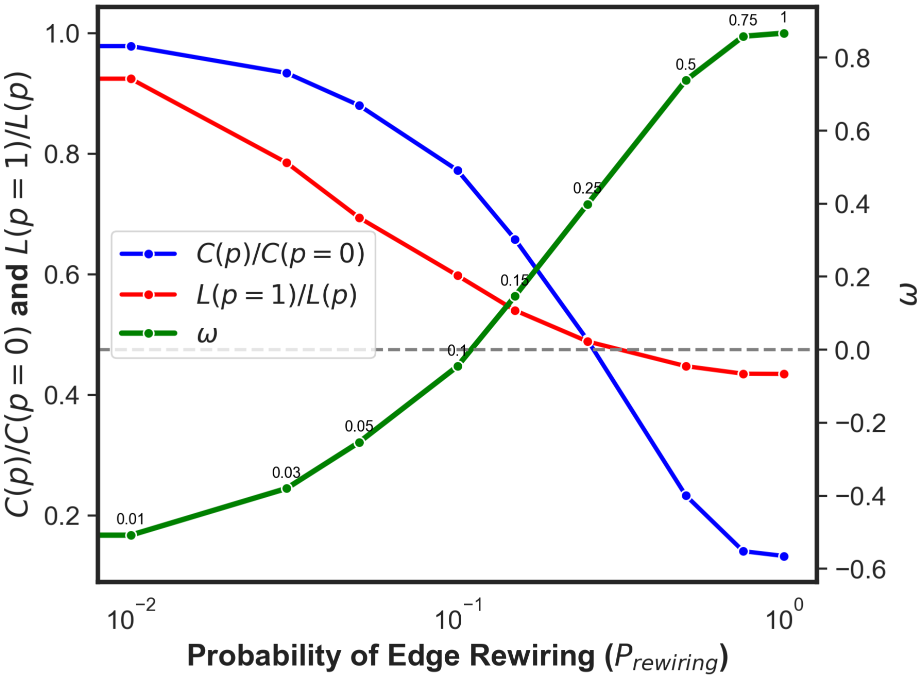
[64] Rajpal, H. & Guerrero, O. Synergistic small worlds that drive technological sophistication. *PNAS Nexus* pgaf102 (2025).

[65] Vinicius, L., Rizzo, L., Battiston, F. & Migliano, A. B. Cultural evolution, social ratcheting and the evolution of human division of labour. *Philosophical Transactions B* **380**, 20230277 (2025).

[66] Rasenberg, M., Özyürek, A., Bögels, S. & Dingemanse, M. The primacy of multimodal alignment in converging on shared symbols for novel referents. *Discourse Processes* **59**, 209–236 (2022).

[67] Falandays, B. J. *et al.* All intelligence is collective intelligence. *Journal of Multiscale Neuroscience* **2**, 169–191 (2023).

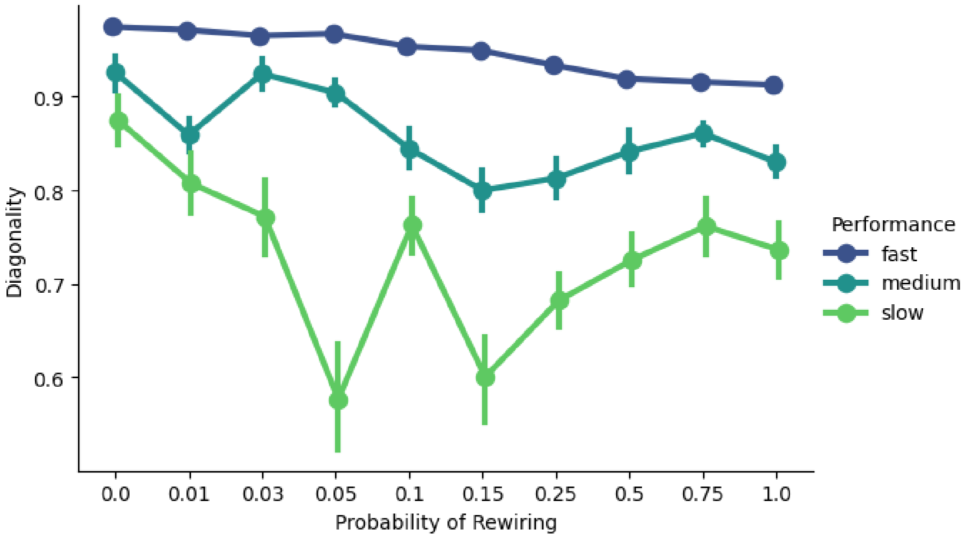
8 Supplementary Figures



Supplementary Figure 1: Small-world index (ω, on the right y-axis) as a function of probability of rewiring (x-axis). ω is calculated as $L(p=1)/L(p) - C(p)/C(p=0)$. $L(p)$ and $C(p)$ refer to the average shortest path length and average clustering coefficient of networks with $P(rewiring) = p$, respectively. $L(p=1)$ is the average shortest path length in an equivalent random graph. $C(p=0)$ is the average clustering coefficient of an equivalent lattice graph. Networks in the small-world regime have ω values closer to 0 (shown by the dashed line).

Diagonality Index : In order to measure how closely each network's trajectory aligns with the 45° diagonal in the problem-space, we represent changes in potion generation for trajectories A and B as a movement vector (a, b), where a and b are the changes in the number of potions generated for A and B, respectively. Summing these stepwise changes yields a net displacement, and its angle θ (computed via $\arctan\left(\frac{b}{a}\right)$) is compared to 45° ($\pi/4$). We then define the Diagonality index as:

$$D.I. = 1 - \frac{\left|\theta - \frac{\pi}{4}\right|}{\frac{\pi}{4}},$$



Supplementary Figure 2: The average Diagonality index of the “movement” of networks with different values of the rewiring parameter over the problem-space, for each performance category. A perfectly diagonal movement will have a Diagonality index of 1.

such that perfectly diagonal movement returns a value of 1, and increasing deviations from 45° approach 0.

For each movement vector, we calculated the angular deviation from an ideal diagonal and normalized them such that perfect diagonal movement along the problem space would be scored 1 and deviations from it would be lower. This index shows that fast networks solve the problem in a more diagonal and optimal fashion than medium and slow networks, irrespective of the $P_{rewiring}$.