PNAS RESEARCH ARTICLE | ECOLOGY SOCIAL SCIENCES OPEN ACCESS

Comparing cooperative geometric puzzle solving in ants versus humans

Tabea Dreyer[a], Amir Haluts[b], Amos Korman[c], Nir Gov[b], Ehud Fonio[a], and Ofer Feinerman[a,1]

Edited by Marcus Feldman, Stanford University, Stanford, CA; received July 17, 2024; accepted November 11, 2024

Biological ensembles use collective intelligence to tackle challenges together, but suboptimal coordination can undermine the effectiveness of group cognition. Testing whether collective cognition exceeds that of the individual is often impractical since different organizational scales tend to face disjoint problems. One exception is the problem of navigating large loads through complex environments and toward a given target. People and ants stand out in their ability to efficiently perform this task not just individually but also as a group. This provides a rare opportunity to empirically compare problem-solving skills and cognitive traits across species and group sizes. Here, we challenge people and ants with the same "piano-movers" load maneuvering puzzle and show that while ants perform more efficiently in larger groups, the opposite is true for humans. We find that although individual ants cannot grasp the global nature of the puzzle, their collective motion translates into emergent cognitive skills. They encode short-term memory in their internally ordered state and this allows for enhanced group performance. People comprehend the puzzle in a way that allows them to explore a reduced search space and, on average, outperform ants. However, when communication is restricted, groups of people resort to the most obvious maneuvers to facilitate consensus. This is reminiscent of ant behavior, and negatively impacts their performance. Our results exemplify how simple minds can easily enjoy scalability while complex brains require extensive communication to cooperate efficiently.

collective intelligence | social insects | human behavior | cooperative transport | consensus decisions

Group living holds many advantages, of which "collective cognition" (1–4) is particularly intriguing. This joint cognition (5) fulfills the basic requirements for cognition according to Shettleworth (6, 7) as it allows the group to sense (8–11), integrate (12–15), and respond (16, 17) to environmental cues. Collectively improving these abilities allows a cooperative group to expand its cognitive capacities (5) beyond those of its individual members. On the other hand, since large ensembles may be difficult to coordinate, their collective efforts are often counterproductive (18–21). While tempting, direct comparisons between the cognitive abilities of wholes versus individuals are often meaningless because different group sizes tend to interact with the environment at different scales. For example, there is no meaningful way to compare the cognitive capacity of a single neuron to that of the brain. Although there are several recognized instances of the emergence of novel cognitive tools on the collective level (8, 22), efforts to directly and quantitatively compare an agent's cognition with that of the group remain quite rare (23–27).

Species that engage in cooperative load transport offer a unique opportunity for biologically meaningful cross-scale comparisons. In these species, small loads are hauled by individuals, while large loads are transported by groups. Naturally, grouping increases muscle power. However, if it also enhances cognitive abilities, this can significantly improve navigation in cluttered environments (10). Quantitatively testing for such improvements can be done by employing physical puzzles (28, 29) and scaling them appropriately to challenge either individuals or groups. In this work, we implement this using the piano-movers puzzle (30, 31)—a maneuvering task in which an oddly shaped load must be transported across a tight and obstructed environment. Being purely geometrical, this puzzle is completely scalable in size. It can be made small enough to be tackled by a single individual or scaled up to sizes that require the cooperative effort of a large group. Such large-scale cooperative transport is extremely rare in nature and performed mainly by people and around 1% of known ant species (3, 32, 33). Here, we study individuals and groups of both people and longhorn crazy ants (*Paratrechina longicornis*) as they tackle this puzzle.

Significance

Collective cognition is often mentioned as one of the advantages of group living. But which factors actually facilitate group smarts? To answer this, we compared how individuals and groups of either ants or people tackle an identical geometrical puzzle. We find that when ants work in groups, their performances rise significantly. Groups of people do not show such improvement and, when their communication is restricted, even display deteriorated performances. What is the source of such differences? An ant's simplicity prevents her from solving the puzzle on her own but facilitates effective cooperation with nest-mates. A single person is cognitively sophisticated and solves the problem efficiently but this leads to interpersonal variation that stands in the way of efficient group performance.

Author affiliations: [a]Department of Physics of Complex Systems, Weizmann Institute of Science, 7610001 Rehovot, Israel; [b]Department of Chemical and Biological Physics, Weizmann Institute of Science, 7610001 Rehovot, Israel; and [c]Department of Computer Science, University of Haifa, 3303221 Haifa, Israel

Author contributions: T.D., E.F., and O.F. designed research; T.D., A.H., N.G., E.F., and O.F. performed research; T.D., A.H., A.K., and O.F. analyzed data; T.D., A.H., A.K., N.G., and O.F. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: ofer.feinerman@weizmann.ac.il.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2414274121/-/DCSupplemental.

Published December 23, 2024.

One Puzzle, a Variety of Solvers

We designed a piano-movers puzzle in which a T-shaped load is to be maneuvered across a rectangular arena divided into three chambers that are connected by two narrow slits (Fig. 1 *A* and *B*). To solve the puzzle, the load must be transported from its initial location in the left-most chamber (*a* in Fig. 1*C*), through the second chamber, and into the third chamber (*h* in Fig. 1*C*), which is open to the outside. An "optimal solution" of the puzzle (without any unnecessary steps) follows the sequence of maneuvers highlighted by green arrows in Fig. 1*C*.

We presented scaled versions of this puzzle to both people and ants (Fig. 1 *A* and *B*). People attempted to solve the puzzle because they were instructed to, while ants were motivated to carry the load to the third chamber (which was open toward the nest) since the load was made to resemble food. The puzzle was designed to pose significant challenges for both species. People are challenged by the precise length assessments, mental rotations, and symmetry comprehension (*SI Appendix*, Fig. S1 *A–C*) (34, 35) that are required to distinguish between viable moves and dead-ends. The puzzle is challenging for ants since their pheromone-based communication takes neither load size versus door size nor load rotations into account (36), and this deems a major part of their collective navigation strategy useless.

To compare performances across species and group sizes, we prepared the puzzle in five different sizes (*SI Appendix*, Fig. S2 *C–H*). Ants were presented with two different scales of this puzzle: the "small" ant puzzle was tackled either by a single ant ($n = 3$ composite experiments; see *Materials and Methods*) or by a small group averaging about 7 ants ($n = 48$), and the "large" ant puzzle was tackled by a large group averaging about 80 ants ($n = 28$). People were presented with three different puzzle sizes: the small human puzzle was tackled by a single person ($n = 61$), the "medium" human puzzle was tackled by a group of 6 to 9 people, and the large human puzzle was tackled by a group of 16 to 26 people. Snapshots from solution attempts by a selection of solvers are shown in Fig. 1 *A* and *B*, and recordings of solution attempts can be seen in Movies S1–S5.

To mirror the transport behavior of ants, the human and ant puzzles maintained a similar ratio of body size to load size.

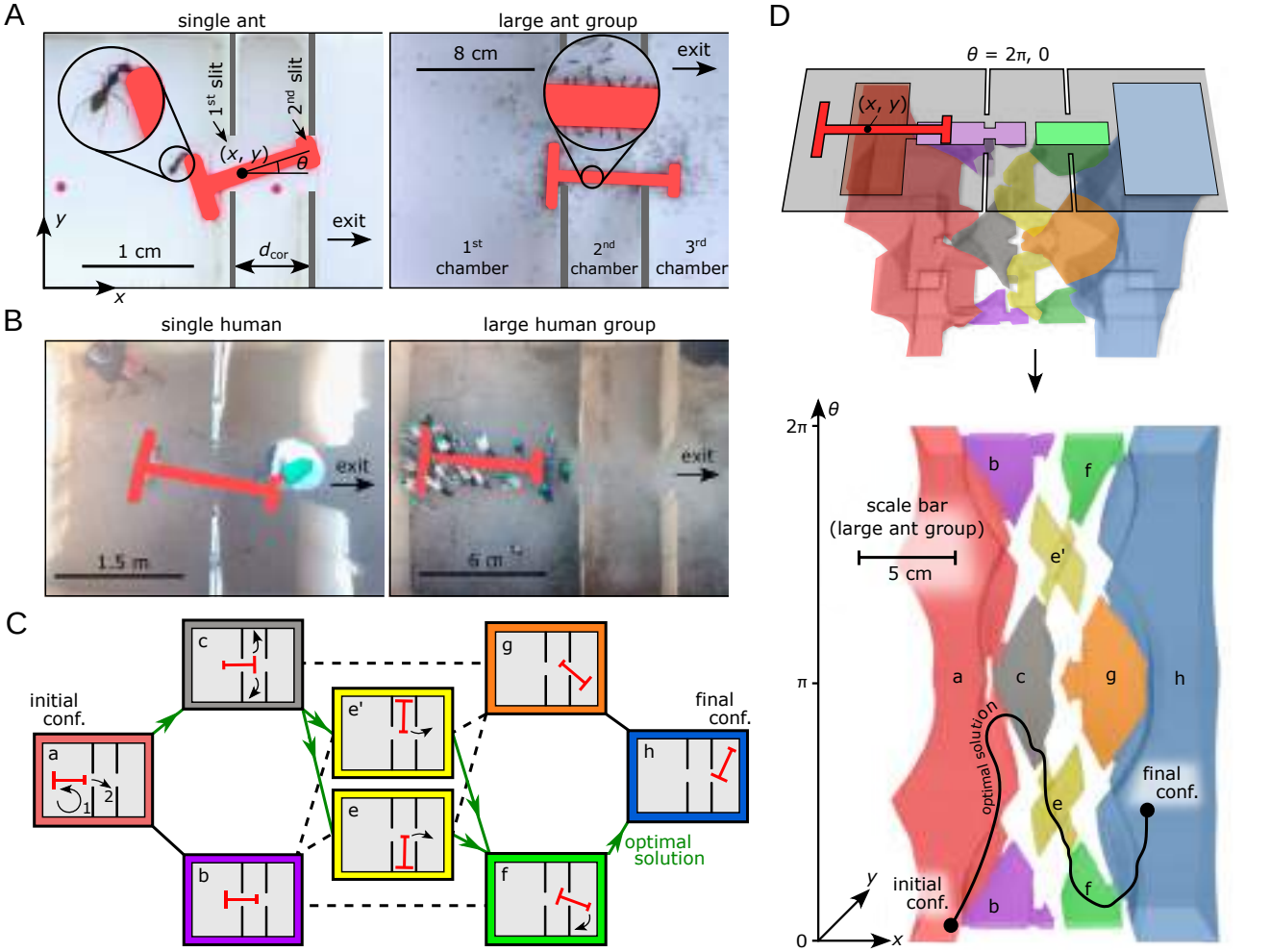

Fig. 1. The piano mover's puzzle. (*A*) Snapshots from two ant solvers: a single ant (*Left*) and a large group of ants (*Right*) during a solving attempt, with zoom-ins on carrier ants. The load's configuration within the puzzle is defined by its center-of-mass position (x, y) and its orientation angle θ. (*B*) Snapshots from two human solvers: a single person (*Left*) and a large group of people (*Right*) during a solving attempt. In panels *A* and *B*, the load in each snapshot is emphasized in red. (*C*) Possible states of the load within the puzzle. Each state is representative of a subvolume in configuration space (panel *D*) and shares a matching color. Solid and dotted lines connect states whose configurations are connected and disconnected (but geometrically adjacent) in configuration space, respectively. The sequence of maneuvers comprising the optimal solution of the puzzle is highlighted with green lines and arrowheads. (*D*) Illustration of the three-dimensional (x, y, θ) configuration space from two different perspectives, where the colors indicate the division into the subvolumes that correspond to the states represented in panel *C*. The top perspective shows a planar slice along $\theta = 2\pi$ (or $\theta = 0$), where the shown load configuration corresponds to state *a*. In the bottom perspective, the black trajectory follows the optimal solution in configuration space.

Further, inspired by ant carriers that maneuver the load by pulling it at their points of attachment (37), each member of a human group was assigned to a handle connected to the load and instructed to exclusively pull on it. The handles of the large load contained force meters, which measured the pulling force applied by every person throughout the attempt. In addition, groups of people were divided into two experimental conditions: unlimited communication (medium: $n = 22$, large: $n = 21$) and highly restricted communication (medium: $n = 28$, large: $n = 20$), where in the latter, the participants were not allowed to speak or use gestures and wore masks and sunglasses to hide their mouths and eyes. Importantly, communication-restricted people groups and ant groups rely predominantly on communication through forces transmitted through the load. For people, this was the main form of allowed communication. For longhorn crazy ants, communication in the context of cooperative transport is naturally mediated by both haptic sensation (38) and pheromone communication (36). However, since in the context of our puzzle, pheromones are practically useless (see above), this primarily leaves the ants with force-based communication. This makes comparisons between ant groups and restricted communication human groups especially compelling.

In what follows, we refer to each of the eight groups mentioned above as a different "solver" (5).

Measuring a Solver's Performance

The kinematics of a solver's attempt can be described by the time evolution of the load's configuration $\mathbf{r}(t)$ in terms of its three degrees of freedom, $\mathbf{r}(t) \equiv (x(t),\ y(t),\ r_{\mathrm{av}}\theta(t))$, where $x(t)$, $y(t)$ are the coordinates of the load's center of mass and $\theta(t)$ is the load's orientation angle (as shown in Fig. 1*A*). The orientation is a cyclic coordinate which is multiplied by the average distance of the load's perimeter from its center of mass, r_{av} (*SI Appendix*, Tables S1 and S2) so that $r_{\mathrm{av}}\theta(t)$ measures the average walking distance of an attached carrier during load rotation. To compare solution attempts across scales, we normalize $\mathbf{r}(t)$ by the second chamber's width d_{cor} (Fig. 1*A*). The puzzle's "configuration space," i.e., the set of all geometrically permissible load configurations is presented in Fig. 1*D*. As illustrated by the black trajectory (Fig. 1 *D*, *Bottom*), the dynamics of any solution attempt can be understood as a continuous trajectory in this space. To analyze solution trajectories, we further split configuration space into smaller volumes that correspond to the "states" depicted in Fig. 1*C*.

As a global measure of each solver's performance, we plot the percentage of attempts that were solved up to a given normalized path length in configuration space (Fig. 2*A*, for both ants and people) and up to a given number of attempted state transitions (Fig 2*B*, for people). We find that, on average, human solvers perform better than ant solvers (Fig. 2*A*). However, the full performance distributions do display a small overlap as the best ant solvers outperform the worst human solvers.

We find that large ant groups perform significantly better than individual ants and small groups of ants. Although our single ant data is more sparse, it is sufficient to demonstrate that a sequence of successive maneuvers by single ants can indeed solve the puzzle and to show that the resulting performance is significantly inferior to that displayed by a large group of ants.

While restricted-communication groups of people communicate via forces as ants do, they did not show a corresponding improvement with numbers. In fact, the opposite was true as these groups performed significantly worse than individuals (Fig. 2*B*). Groups which were allowed to communicate reverse this effect and marginally outperform individuals (Fig. 2*B*).

In the following, we analyze the relevant properties of each solver's search trajectories and force application dynamics as a means of relating its cognitive properties and cooperative strategies to its overall performance.

Ant Solvers

To understand why large ant groups outperform small ones, we first focus on their kinematics within the narrow confines of the puzzle's arena. When a large load, carried by a large group of ants, collides with the boundary, it does not stop or retract (Fig. 3*A*). Rather, the changes to its direction of motion and speed are minimal (*SI Appendix*, Fig. S5 *A* and *B*), and it continues to persistently slide along the boundary for an extended distance (Fig. 3*C*) (39). Conversely, when a small load, carried by a small group of ants, collides with the boundary, it tends to lose much of its speed (Fig. 3*B*), and when motion is resumed, the load

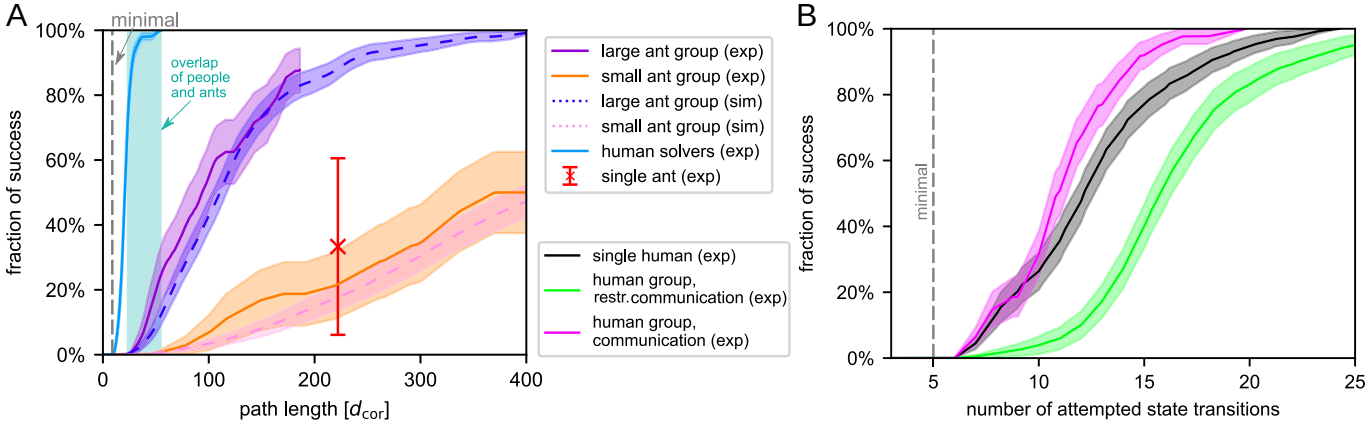

Fig. 2. Performance of different solvers. (*A*) Cumulative distribution function of successful solving attempts as a function of the attempts' path lengths for humans and ants (experiment and simulation). (*B*) Cumulative distribution function of successful attempts as a function of attempted state transitions for single humans and human groups with communication and with restricted communication. Since the performance of medium and large groups is similar in relation to the performance of single humans (*SI Appendix*, Fig. S4*B*), we merged the data of these two group sizes for the sake of clarity. Both cumulative distribution functions were smoothed using a moving average that averaged over a tenth of the maximal measured path length or number of attempted state transitions of the respective solvers. The vertical dotted lines represent the minimal path length and number of transitions required to solve the puzzle in panels *A* and *B*, respectively.

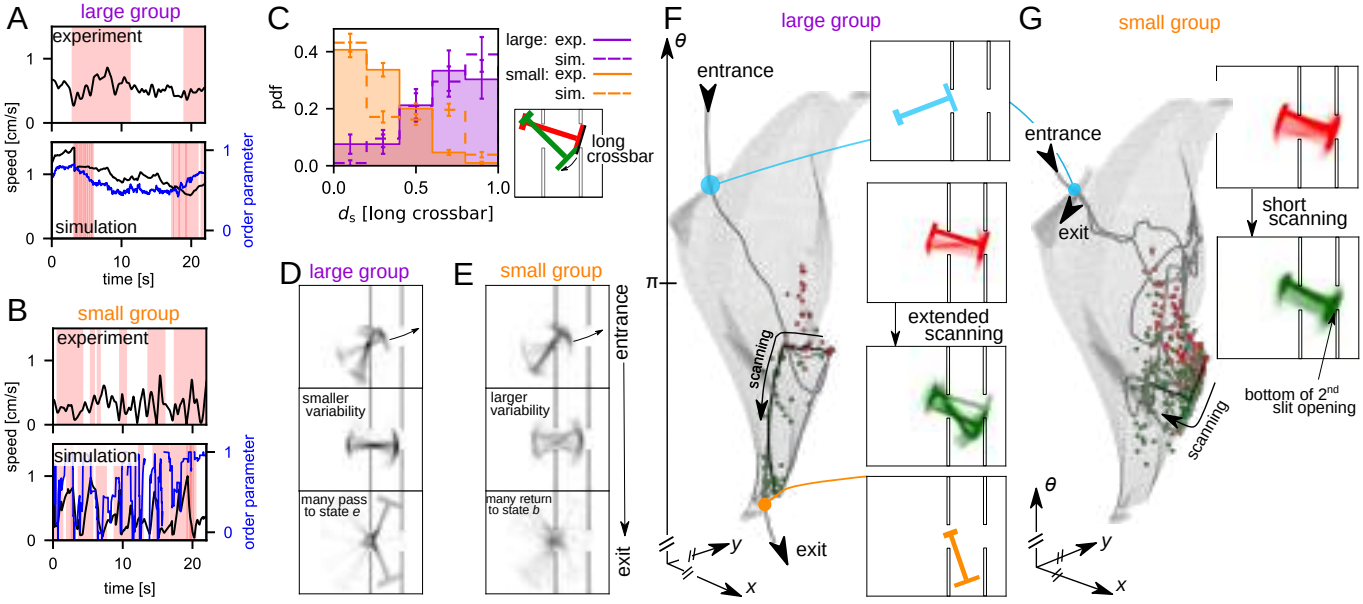

Fig. 3. Motion properties and heuristics of ant solvers. Load's path speed $|\mathbf{v}(t)| = |\dot{\mathbf{r}}(t)|$ over time in the first chamber of (*A*) a large and (*B*) a small group in experiment (*Top*) and simulation (*Bottom*). The red-shaded areas correspond to the load being in contact with the boundary. The blue curves additionally show the order parameter measured in simulations (*Materials and Methods*). (*C*) Probability distribution function of persistent sliding distance, d_S, normalized by the length of the load's long crossbar when it is in contact with the bottom of the second slit opening highlighted in panel *G* (*SI Appendix*, Fig. S5 *C*–*F* and Note 8). (*D* and *E*) Overlaid configurations of the large and small groups, respectively, within state *c* in experiments. The *Top* and *Middle* subpanels show configurations in which, after entering state *c*, the solver reached certain *x*-values for the first time. The *Bottom* panels show configurations at the exit from state *c*, either back to state *a* or to state *e*. While large groups move in a nearly deterministic manner, shown by the smaller variability in configurations, smaller groups exhibit greater variability. (*F* and *G*) Sample trajectories of a large and a small group, respectively, in state *c* in configuration space (only the volume corresponding to this state is shown). The red and green markings correspond to initial and final configurations, respectively, of persistent sliding (scanning) segments along the bottom of the second slit opening (marked in panel *G*). The overlaid red and green configurations plotted on the *Right* in panels *F* and *G* correspond to the red and green markings, respectively. The blue and yellow configurations illustrate the trajectories' entrances and exits from state *c*. For the large group, the load moves persistently along the boundary, which channels it to the narrow passage leading to the next state. This persistent sliding motion, which is not aligned with the exit direction, is absent in small groups.

often takes a direction that is different from the original one (*SI Appendix*, Fig. S5 *A* and *B*). This lack of persistence causes the small load to either quickly retract from the boundary or exhibit much shorter bouts of persistent sliding (Fig. 3*C*). These differences between large groups and small groups align with previous work, which showed that direction changes by a larger group of ants require disproportionately larger perturbations (39, 40). Single ant solvers frequently detach from the load before reattaching to it at a different location. This randomizes the load's direction of motion and prevents individuals from exhibiting extended persistent wall-sliding motion (*SI Appendix*, Fig. S5*G*). Therefore, small groups serve as a good approximation for single ants in both local kinematics and overall performance (Fig. 2*A*).

To understand the dynamics behind these kinematic observations, we employed an agent-based model of ant cooperative transport (37, 38). The model specifies the noisy force application rules at the level of individual ants (*Materials and Methods*). In brief, this empirically verified model assumes that when an ant attaches to the load, she transiently acts as an "informed leader" by pulling it in the direction of the nest. After about ten seconds, the newly attached ant switches her state to that of an "uninformed follower" and tends to align her pulling effort with the direction in which the load, at her point of contact, is moving. This tendency is larger for larger group sizes. We have previously shown that in the absence of boundaries, these rules lead to ballistic motion for large loads and biased random walk for smaller ones (3, 37, 38). We simulated the movement of a load carried by ants that adhere to these rules within the puzzle's confines using a physics engine (41). In these simulations, the load was subjected to forces from individual ants, as well as normal forces applied by the walls during collisions (*Materials and Methods* and Movies S6 and S7).

The simulation provides a dynamic record of the forces each ant exerts before and during a collision. This allows us to measure the degree to which the ants are coordinated and explains the observed differences between the two load sizes. This was done by considering an order parameter defined as the average cosine angle between an ant's pulling direction and the load's direction of motion at her point of attachment (*Materials and Methods*). The dynamics of the order parameter show that large groups maintain strong alignment when they are within confined chambers and that this alignment persists even when they collide with a wall (Fig. 3*A* and *SI Appendix*, Note 2). Intuitively, the fact that the load maintains its motion provides mechanical coupling between the carrying ants, which, in turn, results in high coordination and persistent sliding motion (Fig. 3 *A* and *C*). For small groups, drops in load speed (Fig. 3*B*) prohibit efficient coordination. This causes the ants to lose directional persistence (Fig. 3 *B* and *C*)—an effect that is somewhat more pronounced in experiments than in simulation (Fig. 3*C*), which could be explained by several effects that were excluded from the model for the sake of simplicity (*SI Appendix*, Fig. S3).

Our simulation captures more than just local dynamics. It provides qualitative agreement with experiments in that large groups significantly outperform the small groups (*SI Appendix*, Fig. S4*A*). To account for the fact that the body length of an ant, at approximately 3 mm, does not scale with the puzzle's size, we also ran finite-size corrected simulations in which both small and large loads were expanded by an impenetrable thin rim sized at 10% of an ant's length. The corrected simulations provide

quantitative agreement with empirical measurements of overall puzzle-solving performance for both small and large groups of ants (Fig. 2*A*).

Importantly, the high degree of quantitative agreement between experiment and simulation was achieved while assuming simple ants with basic cognitive traits. Indeed, all simulated ants follow local rules that do not require any large-scale understanding of the puzzle's geometry (beyond the location of the next slit's opening) and that are completely independent of group size or geometrical context. This strengthens the assumption that the improved puzzle-solving abilities of large ant groups are not the result of individual cognition, but rather emerge from interacting individuals who make no geometrical considerations.

Translating the ants' mechanical traits into the language of cognition allows us to describe the cognitive faculties that allow large ant groups to solve the puzzle successfully. The high persistence of large groups translates into short-term collective memory. This emergent memory leads to thigmotactic wall-sliding behavior (Fig. 3*C*). It enables the large group to efficiently scan large areas of the wall without changing their direction (Fig. 3*F*), unlike the small ant group (Fig. 3*G*). They do so until they find an opening and move through it (Bottom panels of Fig. 3 *D* and *E*). More generally, memory allows the large load to move more deterministically (*Middle* panels of Fig. 3 *D* and *E*) than the small one and to exhibit systematic motion that allows for efficient transitions along the different regions of the puzzle (see first two panels in Fig. 4*A* for experimental results and *SI Appendix*, Fig. S6*B* for results from simulation).

To summarize, emergent cognitive faculties allow large ant groups to employ a heuristic that is reminiscent of the well-known "right-hand rule," in which, upon entering a maze, the solver slides their right hand along the wall and proceeds forward without changing their direction. Moreover, the fact that the ants occasionally move away from the wall and collide again at a new location (*SI Appendix*, Fig. S5*I*) allows them to avoid infinite loops that may plague strict right-hand-rule followers (42). On the other hand, small ant groups exhibit random-walk-like dynamics that include futile searches and trapping in dead ends. The superiority of the large group heuristic is in line with previous research on algorithms for maze solving (43–45).

Human Solvers

We have seen that large groups of ants often outperform small ones by following the most direct path when transitioning between states (Fig. 4*A* and *SI Appendix*, Fig. S7*B*). The motion exhibited by people, either as individuals or in groups, displays even higher efficiency as these solvers tend to take the direct path while crossing any configuration space subvolume that corresponds to the states of Fig. 1*D* (*SI Appendix*, Fig. S6*C*). Indeed, a meaningful description of human puzzle-solving behavior can be obtained when considering these states and assuming that people reduce the three-dimensional configuration space to a mental graph-type representation wherein states serve as the nodes (Fig. 1 *C* and *D*). As an example, the node that corresponds to state *a* comprises all configurations in which "the load is fully contained in the first chamber." When at a node, people have no direct means to distinguish which neighboring nodes in configuration space are truly connected to it (34, 35). Therefore, edges represent possible connections between physically adjacent states, irrespective of whether they

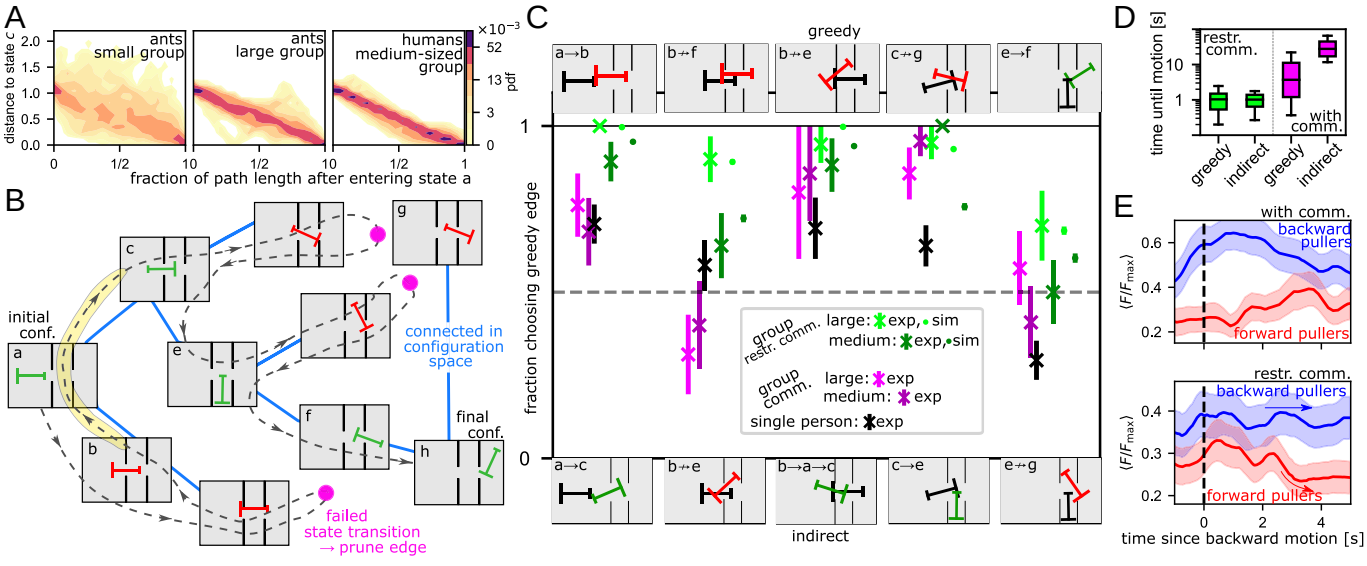

Fig. 4. Motion properties and heuristic of human solvers. (*A*) Two-dimensional heat maps showing the probability of being a certain normalized distance from target state *c*. The probability is shown as a function of the fraction of path length traveled from the initial entry to state *a* after exiting state *b* (see path marked in yellow in panel *B*). Segments of small and large ant groups and medium-sized human groups are shown. (*B*) Schematic of the configurations traversed during a sample single human's attempt illustrating the depth-first-search method. The dotted line shows the order of configurations traversed while succeeding and failing at transitioning between connected and disconnected states, respectively. (*C*) Probability of attempting to traverse the edge leading to the greedy option (*Top*) or the indirect option (*Bottom*) when encountering five common decision forks in the puzzle consisting of two untested edges. Possible edges are depicted by a starting configuration (black load) and ending configurations (green and red loads for edges that are part of or deviations from the optimal solution, respectively) and labeled with *state*1 → *state*2 or *state*1 ↛ *state*2, for existent and nonexistent edges, respectively. Data for different human solvers in experiments and simulations are presented. The simulation's error bars were smaller than the size of the marker. (*D*) Time until first motion of groups of people with and with restricted communication at the start of their attempt. The labels greedy and indirect refer to the top and bottom options shown in the leftmost decision fork in panel *C*. (*E*) Average force of pullers toward and away from the exit direction in red and blue, respectively. To allow for motion against the exit direction ($t = 0$) when retracting from a dead end (e.g., state *b*) and avoid tugs-of-war, pullers toward the exit (red) must release their pull. This occurs before motion begins in communicating groups (*Top*) and after it begins in groups with restricted communication (*Bottom*). The forces were measured using force meters at every participant's attachment location.

are actually connected in configuration space. In Fig. 1*C*, these connected and disconnected edges are represented as solid and dotted lines, respectively.

As people's representation of configuration space possibly contains disconnected edges, they resort to a trial-and-error tactic to find a sequence of nodes that connect the initial to the final node: when in a certain "decision node," people attempt to reach an adjacent "target node" by moving along the shortest path between them in configuration space (Fig. 4*A* and *SI Appendix*, Fig. S6 *A–C*). This demonstrates people's understanding that the shortest path from initial to final configuration consists of shortest paths between connected states in configuration space. If the edge between the decision and target node is found to be missing, they mentally prune it (Fig. 4*B*). This exploration employs long-term memory as people rarely attempt to traverse a previously pruned edge (*SI Appendix*, Fig. S7*A*). Importantly, this is true regardless of the particular location they returned to within the decision node after pruning (*SI Appendix*, Fig. S6*D*). This further supports the hypothesis that people aggregate large components of configuration space (states) into discrete nodes.

While all single-person solvers search the graph in this manner, there is still significant variation in their overall performances (Fig. 2*B*). There are several considerations that people may take into account when deciding which edges to traverse. At node *a*, for example, geometric considerations can affect the decision to pass either the large or small crossbars of the T-load through the first slit. Another consideration is that states e and e' are equivalent (*SI Appendix*, Fig. S1*B*) and that, therefore, only one of them should be tested. Finally, the left-to-right symmetry of the puzzle as a whole (*SI Appendix*, Fig. S1*C*) means that if the edge $b \rightarrow e$ was found to be missing, then the edge $e \rightarrow g$ must be missing as well and can be mentally pruned without actually trying to cross it. Applying such considerations can explain at least some of the variations between different solvers (Fig. 2*B*). Given these variations, we find that $91 \pm 1\%$ of the attempted node transitions observed in people concur with a depth-first search (DFS) (46) heuristic on the tree graph constructed from the original search graph shown in Fig. 1*C* by deleting nonexistent edges and adding dead-end nodes (Fig. 4*B*).

When individuals group to move a load together, the aforementioned variations make the next move nonevident and require consensus-forming mechanisms. These mechanisms can be expected to vary between groups with communication and restricted communication as is, indeed, reflected by their qualitatively distinct decisions. The nature of this distinction becomes evident by focusing on decision forks in the graph and categorizing the possible target nodes into one of two types: The "greedy option" is the node that, if reached, allows the load to extend the farthest in the exit direction (top configurations in Fig. 4*C*). The "indirect option" is the other option (bottom configurations in Fig. 4*C*). We find that, across all five decision forks, groups with restricted communication choose more greedily than individuals and communicating groups (Fig. 4*C*). These greedy decisions lower the overall performances of groups with restricted communication (Fig. 2*B*).

To study why groups with restricted communication tend to act greedily, we focus on the mechanisms they employ to form consensus. We find that, at the beginning of an attempt, loads that are carried by groups with restricted communication commence motion within less than a second (Fig. 4*D*). The individual force-meters show that similar to ants, once motion begins, all group members tend to align their pulling efforts with it (Fig. 4*E*) (47). Further, we find that during the transition between nodes, the load does not change its direction and does not deviate from the shortest path (Fig. 4*A*). These mechanical observations lead us to suggest a one-shot majority vote for groups with restricted communication: at a decision fork, all group members that are not restricted by the walls and can pull toward at least one of the two options (*SI Appendix*, Fig. S7 *B* and *C*) start exerting a force toward their chosen target node. Summing up the force vectors of all initial pullers is equivalent to a majority vote among this leading subgroup (or *oligarchy*, see ref. 48). Once motion commences, all other group members quickly comply with this direction and transport the load to the chosen target node.

This is a social combination model (49, 50) as people do not discuss the options but, rather, collectively choose a course of action that is based on individually formed opinions. To complete the model, one must specify the initial pulling directions applied by members of the leading subgroup. A natural guess would be to sample these directions from the decisions taken by individual solvers at the same decision point. However, this guess is inconsistent with our observations since groups with restricted communication tend to choose greedily even in nodes where the majority of individuals prefer the indirect move (e.g., right-most decision fork in Fig. 4*C*). Clearly, majority-based votes cannot flip group preferences in this fashion. We, therefore, add a greediness parameter p to our model. Members of the initial subgroup pull toward the node they would have pulled to, had they been alone, with probability $(1 - P)$, and to the greedy option with probability P. Fixing the value of the greediness parameter at $P = 0.2$ provides a reasonably good fit for both medium and large groups in nine out of ten decision forks (Fig. 4*C*).

We can hypothesize on the factors that lead people to increase greediness within a group with restricted communication. One reason may be a result of local, limited perception of the puzzle. Recall that the size of the puzzle is scaled to the size of the group. Hence, unlike a person who solves the puzzle alone, each member within a group occupies a small area of the puzzle, which induces a notion of locality. Since people in groups with restricted communication are unable to efficiently share information from different parts of the puzzle, they react mostly to the opening that is closest to them (51) and pull the part of the object they are holding toward it. These local forces are consistent with motion toward the greedy option.

Greediness may also be related to the empirical observation that when debate is prohibited, people tend to reach a consensus quickly (Fig. 4*D*). In this case, in line with the notion of groupthink (21, 52), people tend to forsake their personal opinion and promote a different one, not because they think it is the better option, but because it is the option they believe is most likely to be independently chosen by others. The fact that a single fitting parameter in our model suffices to describe people for both group sizes suggests that small group sizes are sufficient to induce the maximal groupthink effect. It may be the case that this assessment of the majority opinion (53) is, in fact, the minority opinion (see rightmost decision fork in Fig. 4*C*), which is reminiscent of the social phenomenon of pluralistic ignorance (54) typically demonstrated with regard to more complex social issues. Being based on spatial considerations, the greedy option can be expected to be better than another arbitrary simple consensus. Thus, by choosing the greedy option, the group is able not only to reach a fast consensus but also a reasonably good one. Finally, it is important to note that reaching a suboptimal consensus is not detrimental; even if this attempt leads to a dead end, the failure becomes the new common knowledge across the group (55), and this facilitates consensus on the next decision.

Communicating groups reach consensus in a very different manner: at the beginning of a solution attempt, and before any motion commences, they tend to spend tens of seconds conversing (Fig. 4*D*). In contrast to groups with restricted communication, communicating group members adjust their pulling forces with the direction of the desired target node even before motion begins (Fig. 4*E*). They then maintain this agreement until the target node is reached (Fig. 4*A*). The ability to discuss frees the group from the urge to make a single-shot decision and they can, instead, take their time to advocate for less obvious, but more accurate choices toward a joint decision (Fig. 4*D* and *SI Appendix*, Fig. S6*E*). The fact that a joint decision to move to the indirect option takes significantly longer than a decision to move greedily (Fig. 4*D*) strengthens the claim that communication is crucial for forming a consensus around a less obvious choice.

In most decision forks, the decision probabilities exhibited by communicating groups are similar to those of a typical individual solver (Fig. 4*C*). A reasonable mechanism to implement this is that communication is used to elect a random group member and follow her lead toward the option she would have chosen if alone (supported by personal observations; see *SI Appendix*, Note 6). The end result is that communicating groups perform similarly to individuals (Fig. 2*B*). We note that, in the context of a confusing puzzle where, in all decision forks, the majority of individual choices are wrong, this tactic, intuitively adopted by groups, is strictly superior to majority decisions (*SI Appendix*, Note 10). It is interesting that groups of people naturally revert to this behavior. On the other hand, random leader choice or an equivalent tactic is worse than electing the most competent leader as, statistically speaking, a communicating group will attempt to traverse 35% more edges during a solution than its best-performing individual (*SI Appendix*, Note 9). Therefore, communication does not significantly help the group distinguish a competent member from the rest and follow her lead.

Nevertheless, the communicating group does marginally outperform its average group member (Fig. 2*B*). We find that the primary factor for this improvement stems from the communicating groups' ability to specifically escape the dead-end state *b* without exploring all its cul-de-sacs (*SI Appendix*, Note 7). Apparently, this is done as some group members use speech to demonstrate (56) that this is a dead-end before further futile state transitions are attempted. This implies that a large enough minority of participants can use discussion to sway the group from the common knowledge opinion toward a nontrivial shortcut (48, 50).

Comparing Ants and Humans

The physical realization of the piano-movers problem provides us with two rare opportunities. It allows us to compare problem-solving skills and performances across group sizes and down to a single individual and also enables a comparison of collective problem-solving across species. Here, we focus on comparisons between group sizes within each species and between ants and restricted communication groups of humans, where the latter enables a meaningful cross-species comparison.

Large ant groups exhibit emergent persistence, which expands their cognitive toolbox to include short-term memory—a building block of cognition (6, 7): the memory of the current direction of motion is temporarily stored in the collective ordered state of the transporting ants, analogous to ordered spins in statistical mechanics (38). Thus, collective memory is an emergent feature rather than an individual trait. Similar examples in which grouping leads to newly acquired cognitive abilities are rare. One example comes from fish that can expand their sensing range by grouping, which allows the school to respond to environmental gradients that occur over long length scales (8). Another example comes from fish that modulate their responsiveness to risk by modifying the spatial structure of the group (22).

Emergent memory allows groups of ants to perform near-deterministic, persistent scanning of the wall, which potentially leads them through shortest paths in search space. This emergent boundary-following heuristic enables large ant groups to significantly outperform small groups as well as individual ants. Memory, directed motion, systematic search, and the employment of heuristics are all attributes of human solvers. Thus, the expansion of their cognitive toolbox allows large groups of ants to confront the puzzle in ways that resemble human solvers.

Cooperative transport offers the rare opportunity to experimentally analyze human large-scale (>20) cooperation in the context of a natural behavior with a clear common goal. As a natural behavior, this collective task incorporates numerous aspects of collective problem-solving (48–50, 54, 56–58) whose joint effect is unpredictable. Different from ants, people successfully tackle the puzzle as individuals, but grouping raises an obstacle since consensus is required for efficient motion.

Communicating groups of people spend significant time discussing and deciding on their next move and, by this, display similar performance to individuals. When communication is restricted, people completely replace their social-communication (59) debating heuristic with a faster, social combination heuristic. In this case, they tend to act differently from their thought-over opinion (21, 52) and pull toward the lowest common denominator, the greedy option, as would a newly attached informed ant (37). Once the load starts moving, people in restricted communication groups simply align their pull with its motion. This abandonment of their individual cognitive abilities is, once more, reminiscent of the collective ant behavior (37, 38). As such, when tackling the puzzle with restricted communication, large groups of people display deteriorated performance by adopting some ant-like properties (36, 37). This deterioration is lifted if communication is allowed.

A general understanding of collective problem solving by animal groups can benefit from comparing extreme examples: People stand out for individual cognitive abilities while ants excel in cooperation. In this work, we sharpened the comparison by designing experimental conditions that impose similar, force-based communication on both species (restricted communication condition in humans).

While advanced cognitive capabilities have been shown in ants (60), the agreement between empirical measurements and our agent-based model implies that within our puzzle, individual ants do not employ any large-scale geometrical consideration (see also ref. 36). Therefore, we assume that while longhorn crazy ants discern the context of cooperative transport, they make no distinctions regarding the geometry of the specific problem and always apply the same individual scale behavioral rules (37).

This uniformity ensures that, within a given context, all participating ants operate within the same framework, and this can facilitate cooperation (61) and efficient scaling of cognitive capacities. Such scaling is crucial given the inherent cognitive limitations of a single ant (36). Importantly, the ants' joint efforts lead to an emergent heuristic of wall sliding, which for large enough groups, proves to be highly robust and effective for an extremely wide range of complex environments (10, 36, 39, 40, 62).

People are more flexible in selecting tools from their cognitive repertoire and can finely adjust their problem-solving tactics to suit the particular task at hand (63–66). While this flexibility can enhance individual performance, it inevitably results in interpersonal differences (e.g., ref. 67) that may require more advanced communication to avoid worsening collective performances (Fig. 2*B*) and allow for effective cooperation (68, 69).

These differences between ants and humans illuminate two evolutionary trajectories that differ in the way cognitive abilities are allocated between the individual and collective levels. The comparison between them suggests several research directions: theoretical study is required to establish the trade-off between complexity at the individual scale and the ease of cooperation at the group level (see ref. 70 for a similar effort). Comparative studies across phylogenetic trees can be used to explore empirical examples of this trade-off. The lessons learned may be useful in the design of cooperative robotic groups (see, for example, ref. 71) and, perhaps, help us understand why of all the social animals, only ants and people excel at cooperative transport.

Materials and Methods

Setup for Experiments with Ants. The small ant arena was built by 3D printing the floor and boundaries of the arena using white polylactic acid (PLA). Additionally, a transparent cover was placed on top of the arena that prevented ants from carrying the load over the arena boundary. Between the two slits, another transparent layer was installed to prevent ants from slanting the load while pulling it through the slits and, thereby, "cheating." A schematic of the resulting arena is shown in *SI Appendix*, Fig. S2*A*, and two photographs of a single ant and a small group of ants during a solving attempt are shown in *SI Appendix*, Fig. S2 *C* and *D*, respectively. The large ant arena was built by first laser-cutting rectangular pieces of plastic with appropriate dimensions to form the floor and boundaries of the arena. Then, these pieces were glued together, which constituted the large ant arena. The arena contained holes opening to a room left of the first chamber, where a recruitment ring could be placed. The ring could be used to increase the number of ants in the arena during the experiment to sustain cooperative transport of the load. A schematic of the resulting arena is shown in *SI Appendix*, Fig. S2*B*, and a photograph of a large group of ants during a solving attempt is shown in *SI Appendix*, Fig. S2*E*.

The loads appropriately sized for the small and large arenas were 3D printed with red polylactic acid (PLA), rendering thin, rigid bodies. These are displayed in their respective arenas in *SI Appendix*, Fig. S2 *C*–*E*. The load of the large arena consisted only of a T-shaped rim in order to preserve a similar circumference-to-weight ratio to the load of the small arena, as the number of ants attached scales roughly linearly with the circumference. The measurements of all resulting arenas and loads are listed in *SI Appendix*, Table S1.

Before the experiments, the boundaries of the arenas were covered with Fluon to prevent ants from escaping over the boundary. We incubated the loads in cat food overnight and rubbed canned tuna on them, which made them seem like attractive food items to the ants.

Procedures during Experiments with Ants. To conduct experiments with either single ants or small groups, we positioned a small arena next to a nest of *P. longicornis* ants at the Weizmann Institute of Science in Rehovot, Israel. For experiments involving large ant groups, a larger arena was used instead. For all group sizes, the respective arena was oriented such that its *x*-axis, as defined in *SI Appendix*, Fig. S1*A*, pointed toward the nest entrance. We mounted a Canon EOS 550D camera above the arena using a rolling stand, as shown in *SI Appendix*, Fig. S2 *A* and *B*. The camera was set to record 50 frames per second at a resolution of 1280×720 pixels. We then recruited ants into the arena by placing a recruitment ring rubbed with canned tuna in the first chamber of the arena. Once enough ants were present in the arena and they began cooperatively transporting the recruitment ring, we removed the ring. Then, we started recording the arena with the camera and presented a T-shaped load to the ants in the first chamber in the initial configuration illustrated in *SI Appendix*, Fig. S2*B*. Ants were able to attach to the thin rim of the load and transport it, as shown in *SI Appendix*, Fig. S2 *C*–*E*.

During single ant experiments, an ant was allowed to enter the arena if there was no other ant already present. If there was, the experimenter blew the newly arriving ant away using his/her breath. During experiments with small and large groups of ants, ants were not restricted from entering the arena. In all experiments, ants were allowed to exit the arena at any time. Possible reasons for exiting the arena included loss of interest in the load and recruitment of more ants to the load. The flux of newly attaching ants in experiments with small and large groups replaced the detaching ants, generally maintaining a sufficient number of ants to sustain cooperative transport. The numbers of ants involved in carrying are given in *SI Appendix*, Table S1. If the flow of ants to the load in the large arena was too low to sustain cooperative transport, we placed the recruitment ring in the room left of the first chamber to increase the number of ants in the arena. An attempt was considered successful once the load reached a configuration where it was entirely within the third chamber, which is labeled in *SI Appendix*, Fig. S2*B*. If the puzzle was not solved by a small or large group within approximately 30 min, the experiment was terminated and considered an "unsuccessful attempt." It proved difficult for a succession of single ants in the arena to carry the load for a significant distance within 30 min. Therefore, we did not terminate the experiments with single ants after 30 min, and instead allowed carrying for multiple hours, during much of which there was no ant present in the arena. Further, two of the three single ant experiments spanned multiple days, where the load was reintroduced in the state that the attempt was last terminated in order to reach longer path lengths. For this reason, we call them "composite experiments." The sparsity of the data is due to single ants commonly leaving the setup in order to recruit further ants. This is very much in line with our results that this species has evolved to greatly benefit from its collective efforts.

Exemplary recordings of a succession of single ants, and small and large groups of ants attempting to solve the puzzle can be seen in Movies S1–S3. Experiments were conducted using five different colonies over the course of 3 y.

Setup for Experiments with People. The three arenas for people, whose sizes we refer to as small, medium, and large, were built at the Weizmann Institute of Science in Rehovot, Israel. The boundaries exceeded 2 m in height and were constructed using metal grids overlaid with tarps to prevent participants from seeing through the boundaries. The load of every arena consisted of a T-shaped plastic tube on a cart with wheels. The tubes had pink strips of tape glued to them for the sake of easier tracking of the load (see Fig. 3*B* and *SI Appendix*, Fig. S2 *F*–*H*). The dimensions of the arenas and loads are listed in *SI Appendix*, Table S1. The arenas had gravel floors, which were flattened to ensure easy rolling of the carts. Loads of the medium and large arenas had handles made of skinny, rigid handles. These were connected to fixed locations around the T-shaped tubes using webbing strap. These handles served as attachment sites for participants from which they were able to pull the load. For some experiments with groups of people, force meters that were integrated into the handles were used to measure the pulling force applied by every participant throughout the attempt. The force meters were taken from devices originally designed to measure a travel suitcase's weight. They were integrated into a custom-built data acquisition system that wrote the measured force at a rate of *m* 2 Hz onto a hard drive. Cameras were installed above all three arenas to record the experiments. The cameras recorded the attempt at a frame rate of 30 frames per second, and snapshots of these recordings are shown in Fig. 3*B* and *SI Appendix*, Fig. S2 *F*–*H*.

Procedures during Experiments with People. To conduct experiments with either a single person or a medium-sized or large group of people, a guide was introduced to the participants. Participants included children ages 10 and above and adults who responded to an advertisement at the Weizmann Institute of Science, and other various groups from across Israel. Every participant wore a bright green hat during the solving attempt. Participants of groups with restricted communication additionally received face masks and sunglasses to prevent communication through lip and eye movement. The guide brought the participant(s) to the arena, where the load was presented in the initial configuration in the first chamber. The guide explained that the load had to be moved to the third chamber. Single people solving the puzzle in the small arena were allowed to pull and push the load. Participants in a group in the medium

and large arena were only allowed to influence the load motion by pulling from a handle assigned to them. For groups with restricted communication, verbal and gestural communication was forbidden. For further details on the exact instructions given, see *SI Appendix*, Note 4.

Once all participants' questions were answered, the guide signaled to the group that they could start moving the load. During the attempt, the guide only interfered when participants violated one of the mentioned rules by reminding them of the relevant rule. The solving process was filmed from the top. All attempts by people were successful, and none exceeded 30 min. Exemplary recordings of a single person and a large group of people who were not allowed to communicate solving the puzzle can be seen in Movies S4 and S5, respectively. Experiments included a total of 1,251 participants. We received permission to work with humans from the Weizmann Institute of Science and the European Union. Ethical approvals for human experiments were given by the Weizmann Institutional Review Board (IRB-Education), by the Israel Ministry of Education Chief Scientist, and by the Ethics Unit of the European Research Council. Experiments were conducted in accordance with their guidelines. Informed consent was obtained from all human participants before the experiments.

Video Analysis. From the recordings of experiments with ants, the center of mass of the load's position x, y and orientation θ was extracted in every frame using a custom-written MATLAB program (MathWorks; version R2019b). The program was based on detecting the red load's pixels by thresholding. It then determined the load's configuration by matching the detected pixels to a template in the form of the load's shape.

From the recordings of experiments with people, another custom-written MATLAB (MathWorks) program was used to extract the center of mass of the load's position x, y and orientation θ from every movie frame. This program first removed the camera distortions from a given frame. Then, using thresholding, the green and pink areas that corresponded to the participants' hats and load markings were found, respectively. Finally, a k-means algorithm was used to find the load's configuration (x, y, θ) in every frame.

The load tracking results were manually corrected if necessary. Further, in order to reduce any small jitters, we smoothed the tracked coordinates in the following manner: we applied a median filter to every coordinate x, y, and θ using a window corresponding to $\Delta t = 1$ s and 0.25 s for single and groups of ants, respectively, and $\Delta t = 2$ s for humans. Within these time frames, no significant velocity changes are expected. Subsequently, we applied a Gaussian filter with window size corresponding to $\Delta t/5$. This effectively eliminated jitters while maintaining the authentic motion of the load.

Simulations of Ant Solvers. To gain further mechanistic insight into the performance gaps between large and small ant groups, we ran simulations of ant solvers based on an established statistical physics model for ant cooperative transport (3, 37, 38). Movies S6 and S7 show animations of simulations illustrating a small and a large group of ants solving the puzzle. In this empirically motivated model, attachment and detachment dynamics of ants establish a cohort of carriers on the load, and their combined pulling forces–which they apply along their body axes–govern the load's motion. Newly attached carriers are transiently informed about the direction of a stationary target (the nearest chamber exit in our puzzle setup) and introduce deterministic directional information into the carried load by directing their pulling efforts toward this target. Once she forgets her initial knowledge of the target's direction, each carrier stochastically switches between the roles of an uninformed puller and a lifter until she detaches. In this role-switching and in the alignment of pullers, uninformed carriers are mechanically coupled through the rigid load. This coupling leads to an uninformed carrier being more likely to become a puller the more she is aligned with the load's local direction of motion at her point of attachment. In addition, uniformed pullers align their pulling efforts with this local motion. These effects lead to emergent coherent motion of the carried load, which is repeatedly steered toward the target due to the transient bits of directional information introduced by informed carriers. More details on the simulations are given in *SI Appendix*, Note 1.

Generalized Order Parameter. While in the previous studies (3, 37, 38), order was synonymous with linear persistence, a solver's performance in our puzzle relies on both linear and rotational persistence. To measure order in our ant solver simulations in a way that accounts for persistence in both translation and rotation, we define an order parameter M, at time t, as

$$M(t) = \frac{1}{N_{\mathrm{p}}(t)} \sum_i n_i^p \, \mathbf{p}_i(t) \cdot \hat{\mathbf{v}}^i_{\mathrm{loc}}(t), \qquad [1]$$

where N_{p} is the total number of pullers (informed and uninformed), $\mathbf{p}_i$ is the direction of force applied by the puller at attachment site i, $\hat{\mathbf{v}}^i_{\mathrm{loc}}$ is the direction of local velocity sensed at site i, and $n_i^p = 1$ if site i is occupied by a puller ($n_i^p = 0$ otherwise). For more details on the simulations, see *SI Appendix*, Note 1. Since M gauges the alignment of pullers with local velocities on the edges of the load, it can capture not only the alignment of pulling forces along the direction of load translation but also the order in puller arrangements that promote persistent load rotation.

Path Length. The path length traversed during a given time period is defined as the cumulative distance traversed in configuration space. This is a proxy for the average path walked by a participant connected to the load. In order to compare path lengths between solvers moving within the puzzle arenas at different scales, we normalize the path length by a scaling factor, namely the distance between the two slits, d_{cor}, marked in Fig. 1*A*, calculated by

$$\frac{1}{d_{\mathrm{cor}}} \sum_{n=1}^{N} \sqrt{(x_{n+1} - x_n)^2 + (y_{n+1} - y_n)^2 + r_{\mathrm{av}}^2(\theta_{n+1} - \theta_n)^2}, \qquad [2]$$

where (x_n, y_n) is the load's center-of-mass position, θ_n is the load's orientation angle at frame (or time step) n and N is the total number of frames (or time steps).

The cumulative change in orientation was multiplied with the average distance of the load's periphery to the center of mass r_{av}, which scales linearly with the load and arena size. Given that ants and people move with the load while attached to the periphery, the length $r_{\mathrm{av}}\theta(t)$ meaningfully takes into account the distance walked by an average carrier when the load rotates, as the load rotates by an angle θ around its center of mass. We numerically calculated r_{av} by averaging over the distance of equally spaced points around the load's periphery to the center of mass.

The Load's Path Speed. The load's path velocity at a given time corresponding to index n is defined as

$$\mathbf{v}_n = \frac{1}{t_{n+1} - t_n} \begin{pmatrix} x_{n+1} - x_n \\ y_{n+1} - y_n \\ r_{\mathrm{av}}(\theta_{n+1} - \theta_n) \end{pmatrix}. \qquad [3]$$

The load's path speed is the Euclidean norm of the path velocity, $v_n = ||\mathbf{v}_n||$.

States and Transition Regions. Configuration space is split into states as illustrated in Fig. 1 *C* and *D*. Every permissible load configuration is part of a state, while some load configurations are additionally part of so-called transition regions, which are defined in *SI Appenidx*, Table S3 and illustrated in *SI Appendix*, Fig. S1 *D* and *E*. An attempt to move to an adjacent yet disconnected state was assumed when participants moved to the transition region within their current state that was close to this target state.

We differentiate between transition regions b_{f1} and b_{f2}, shown in *SI Appendix*, Fig. S1*E*, since we frequently observe people distinctly trying to pass from state $b \rightarrow f$ from two symmetric subregions within the region in state b adjacent to state f. We, therefore, split this region into two separate regions, namely b_{f1} and b_{f2}, in order to penalize both unsuccessful edge traversal attempts (*SI Appendix*, Table S3).

Number of Attempted State Transitions. We assign a state, and potentially a transition region, to every configuration of a search trajectory using the conditions defined in *SI Appendix*, Table S3. We then sum up the number of successful and unsuccessful state transitions in the resulting list of states and transition

regions. Successful and unsuccessful transitions are identified by counting the number of stays in states and transition regions, respectively, that last longer than one second. This process provides us with the number of attempted edge traversals of the trajectory. An exemplary succession of attempted edge traversals is illustrated in Fig. 4*B*.

Cumulative Distribution Function of a Set of Attempts. To generate the cumulative distribution functions for sets of solving attempts, as shown in Fig. 2 and *SI Appendix*, Fig. S4, we initially compute the desired performance measure for every individual attempt of the set. This performance measure r is either the cumulative path length during the attempt (Eq. **2**), or the number of attempted edge traversals. In both cases, a larger value of r in a specific experiment corresponds to poorer solving performance.

We find the largest value $r_{\max}$ of the subset of successful attempts, and exclude all unsuccessful experiments that have a value smaller than the $r_{\max}$ from the set of attempts. This renders a set of n' attempts. Then, we determine 100 values, r_i, which are equally spaced between 0 and $r_{\max}$. For every r_i, we divide the number of successful attempts that have better performance, meaning $r < r_i$, by n' rendering the fraction of success p_i at r_i. We then calculate the error according to $\sqrt{\frac{p_i(1-p_i)}{n'}}$. These values ($r_i$, p_i) and their corresponding error, constitute the cumulative distribution function of a set of attempts.

Discretized Configuration Space. Configuration space refers to the set of all configurations $\mathbf{r} = (x, y, \theta)$ in which the load does not overlap with the arena boundary. We compute this space numerically on a regular grid of space points lying inside the following boundaries: $x \in (0, \text{slit2} + \text{load length})$, $y \in (0, \text{arena breadth})$ and $\theta \in (0, 2\pi)$. The lengths are illustrated in the arena's geometry in *SI Appendix*, Fig. S1*A*. For every volume element, we tested whether the corresponding configuration is permissible, resulting in a binary array, which we call discretized configuration space. The grid sizes of the spaces for every puzzle are listed in *SI Appendix*, Tables S1 and S2. They were chosen such that they resolve all connections between large components in configuration space.

Distance from the Boundary, and Criteria for "Contact". We calculated the Euclidean distance from the boundary for every element in the discretized configuration space using the distance_transform_edt function from the Python package scipy.ndimage. With this map of array elements (configurations) and their respective distances to the boundary, we could determine the distance along a trajectory segment. This process was carried out for the histogram depicted in *SI Appendix*, Fig. S5*H*.

For data based on experiments shown in Fig. 3 *A* and *B* and *SI Appendix*, Fig. S5 *A*, *B*, and *I*, we tracked contact manually by watching the recordings. Unless noted otherwise, all other contact detection was done in the following manner: when the trajectory was in a configuration whose corresponding array element had a distance smaller than a set threshold d_{contact}, the load was considered "in contact" with the boundary and vice versa. For data based on simulations shown in Fig. 3 *A* and *B* and *SI Appendix*, Fig. S5 *A* and *B*, we used $d_{\text{contact}} = 0.05$ mm.

Shortest Path between Nodes. We quantified how closely a solver adhered to the shortest path between states by analyzing all trajectory segments of small groups of ants and medium-sized groups of people in which the solver moved from state b to state a and entered state c (Fig. 1 *C* and *D*). First, we transformed these trajectory segments within state a into a list of 100 equally spaced coordinates $\mathbf{r}_N = (\mathbf{r}_1, \ldots \mathbf{r}_{100})$ along the trajectory using the scipy.interpolate.interp1d function. The configuration $\mathbf{r}_1$ is the first configuration after the load entered state a having exited state b; the configuration $\mathbf{r}_{100}$ is the last configuration before the load exited states a and entered state c.

Then, using the skfmm.distance function, we calculated the distance in configuration space of all $\mathbf{r}_N$ to $\mathbf{r}_{100}$ and divided by the initial distance between the starting and ending coordinates. The distance from c at a given fraction $N/100$ of path length after entering state a from state b is then $\mathbf{r}_N^{\text{norm}} = \frac{\|\mathbf{r}_N - \mathbf{r}_{100}\|}{\|\mathbf{r}_1 - \mathbf{r}_{100}\|}$. Heat maps based on these data are shown in Fig. 4*A* and *SI Appendix*, Fig. S6*B*.

Depth-First Search. We assigned a depth to every state and transition region, which form the nodes of our graph (Fig. 4*B*), which quantifies the minimal number of transitions between states and transition regions necessary to reach the initial node. Transition regions (*SI Appendix*, Fig. S1 *D* and *E*) are dead-ends of the graph. For every trajectory of human solvers, we find the fraction of transitions that concur with a depth-first search. Each transition consists of a starting and target node. After entering a new state, the starting node is the former state, and the target node is the current state. After entering a transition region, the starting node is the current state, and the target node is the adjacent state, denoted in the subscript. A transition concurs with a depth-first search if 1) the depth of the target node is equal to or larger than that of the starting node, and the target node has not been visited previously, or 2) since the last visit to a dead-end, the target nodes have consistently had smaller depths than the respective starting nodes. Symmetric regions, such as b_e and b'_e, constitute different nodes, meaning that the solver can visit both regions without violating the depth-first search heuristic. Further, under these rules, we assume that solvers can exclude edges beforehand and, therefore, not visit them without violating the depth-first search. For example, a solver can move through nodes $a \to b \to b_e \to b \to a$ without visiting nodes b_{f2}, b_{f1} and b'_e without violating the depth-first search, as we assume that they previously excluded them. The two primary reasons for people to deviate from this heuristic were, first, reentering the same transition region in an attempt to traverse a nonexistent edge and, second, failing to traverse an existent edge (e.g., $f \to h$), resulting in a return to previously pruned edges.

Decision Forks. A decision fork is defined by, first, a starting condition, which is either the initial configuration in state a or a previous successfully or unsuccessfully attempted edge traversal, and, second, two yet untested edges: the "greedy" and "indirect" edge. The five most common decision forks highlighted in *SI Appendix*, Fig. S7 *B* and *C* and Fig. 4*C* are

(1) At the beginning of the attempt in state a: greedy edge $a \to b$ or indirect edge $a \to c$,
(2) After passing edge $a \to b$: greedy edge $b \to f$ or indirect edge $b \to e$.
(3) After attempting to pass edge $b \to f$: greedy edge $b \not\to e$ or indirect edge $a \to c$.
(4) After passing from edge $a \to c$: greedy edge $c \not\to g$ or indirect edge $c \to e$.
(5) After passing from edge $c \to e$: greedy edge $e \to f$ or indirect edge $e \not\to g$.

We treat symmetric edges (for example $b \not\to e$ and $b' \not\to e'$) as equivalent. Note that at decision fork (5), the solver could also have chosen first to attempt passing edge $e \not\to b$, before edges $e \to f$ or $e \not\to g$. However, this never occurred, as it would have led the solver against the exit direction. Therefore, we do not include this option.

Simulations of Human Decisions at Decision Forks. To analyze how the opinions of individuals in a group with restricted communication translate to a common decision, we employed a simple agent-based simulation that involved the following steps: first, opinions, either aligning with the greedy opinion (opinion 1) or indirect opinion (opinion 0), were assigned to N agents (medium group: $N = 8$, large group: $N = 21$), according to the probability of a single person at each of the five decision forks (Fig. 4*C*). Then, with probability p, every agent switched to opinion 1, though previously they held opinion 0. We then determined the group's decisions as follows: at decision forks (1) and (3), the entire group performs a majority vote (*SI Appendix*, Fig. S7*B*). At decision forks (2), (4), and (5), the majority vote only occurs among a randomly chosen quarter of all agents, as a fraction of the group primarily determined the next attempted edge traversal (*SI Appendix*, Fig. S7*C*). If, in the majority vote, there were an equal number of participants voting for opinions 0 and 1, the group decided for opinion 0 or 1 with equal probability. We repeated the simulation 2,000 times and averaged over all the groups' decisions. By adjusting only parameter P, we find that setting $P = 0.2$ yields good correspondence between the experimental results (Fig. 4*E*), confirming our model choice.

Statistical Information. Solving attempt repeat numbers, n, are given in *SI Appendix*, Tables S1 (experiments) and S2 (simulations). The force measurement plots shown in Fig. 4*E* were based on 8 and 5 experiments in which force meters were functioning reliably for the groups with communication or restricted communication, respectively. All other plots incorporate measurements from all experiments and simulations unless otherwise stated. In all plots, the mean served as the measure of central tendency, and error bars reflected ± 1 SE of the mean unless otherwise noted.

Data, Materials, and Software Availability. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank Matan Feinerman for puzzle design, Michael Oren for initial experiments, Aviram Gelblum for his tracking software and support, and Lena Demmeler for her help in data collection. We thank Guy Han, Yuri Burnishev, Rostyslav Baron, Ami Maimon, Meir Alon, Yeruham Shimoni, and Agza Elihaho for technical support in arena and load construction on all relevant scales. This research has received funding by the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 770964). A.K. was supported by the Israel Science Foundation (https://www.isf.org.il) Grant 1574/24. O. F. is the incumbent of the Henry J Leir Professorial Chair. E. F. is the incumbent of the Tom Beck Research Fellow Chair. N. G. is incumbent of the Lee and William Abramowitz Professorial Chair of Biophysics.

1. A. Dornhaus, N. R. Franks, Individual and collective cognition in ants and other insects (Hymenoptera: Formicidae). *Myrmecol. News* **11**, 215–226 (2008).
2. I. D. Couzin, Collective cognition in animal groups. *Trends Cogn. Sci.* **13**, 36–43 (2009).
3. O. Feinerman, A. Korman, Individual versus collective cognition in social insects. *J. Exp. Biol.* **220**, 73–82 (2017).
4. I. Couzin, Collective minds. *Nature* **445**, 715 (2007).
5. J. R. Larson Jr, C. Christensen, Groups as problem-solving units: Toward a new meaning of social cognition. *Br. J. Soc. Psychol.* **32**, 5–30 (1993).
6. S. J. Shettleworth, *Cognition, Evolution, and Behavior* (Oxford University Press, 2009).
7. U. Neisser, *Cognitive Psychology, Classic Edition* (Psychology Press, 2014).
8. A. Berdahl, C. J. Torney, C. C. Ioannou, J. J. Faria, I. D. Couzin, Emergent sensing of complex environments by mobile animal groups. *Science* **339**, 574–576 (2013).
9. VH Sridhar, F Weissing, I Couzin, "Collective sensing: How animal groups scale noisy gradients.," PhD thesis, Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands (2016).
10. A. Gelblum, E. Fonio, Y. Rodeh, A. Korman, O. Feinerman, Ant collective cognition allows for efficient navigation through disordered environments. *eLife* **9**, e55195 (2020).
11. F. Pedraja, N. B. Sawtell, Collective sensing in electric fish. *Nature* **628**, 139–144 (2024).
12. E. J. Robinson, O. Feinerman, N. R. Franks, How collective comparisons emerge without individual comparisons of the options. *Proc. R. Soc. B: Biol. Sci.* **281**, 20140737 (2014).
13. E. E. Greenwald, L. Baltiansky, O. Feinerman, Individual crop loads provide local control for collective food intake in ant colonies. *eLife* **7**, e31730 (2018).
14. T. Kameda, W. Toyokawa, R. S. Tindale, Information aggregation and collective intelligence beyond the wisdom of crowds. *Nat. Rev. Psychol.* **1**, 345–357 (2022).
15. I. D. Couzin, J. Krause, N. R. Franks, S. A. Levin, Effective leadership and decision-making in animal groups on the move. *Nature* **433**, 513–516 (2005).
16. C. Detrain, J. L. Deneubourg, J. M. Pasteels, "Decision-making in foraging by social insects" in *Information Processing in Social Insects*, C. Detrain, J. L. Deneubourg, J. M. Pasteels, Eds. (Springer, 1999), pp. 331–354.
17. P. K. Visscher, S. Camazine, Collective decisions and cognition in bees. *Nature* **397**, 400 (1999).
18. P. Bescherer, *The Walking Dead and Breakdown Theories of Collective Action* (Springer, 2019), pp. 119–134.
19. L. Soomaroo, V. Murray, Disasters at mass gatherings: Lessons from history. *PLoS Curr.* **4**, RRN1301 (2012).
20. B. A. Nijstad, W. Stroebe, H. F. Lodewijkx, The illusion of group productivity: A reduction of failures explanation. *Eur. J. Soc. Psychol.* **36**, 31–48 (2006).
21. K. Mullane, B. A. Ritter, *Groupthink* (SAGE Publications, Inc., 2023).
22. M. M. Sosna *et al.*, Individual and collective encoding of risk in animal groups. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20556–20561 (2019).
23. C. R. Reid, D. J. Sumpter, M. Beekman, Optimisation in a natural system: Argentine ants solve the towers of Hanoi. *J. Exp. Biol.* **214**, 50–58 (2011).
24. R. Otake, K. Matsuura, Emergence of collective intelligence through behavioral excitatory transmission in maze-solving termites. *SSRN Electron. J.*, 10.2139/ssrn.4106844 (2024).
25. P. McMillen, M. Levin, Collective intelligence: A unifying concept for integrating biology across scales and substrates. *Commun. Biol.* **7**, 378 (2024).
26. D. Biro, D. J. Sumpter, J. Meade, T. Guilford, From compromise to leadership in pigeon homing. *Curr. Biol.* **16**, 2123–2128 (2006).
27. T. Sasaki, B. Granovskiy, R. P. Mann, D. J. Sumpter, S. C. Pratt, Ant colonies outperform individuals when a sensory discrimination task is difficult but not when it is easy. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13769–13773 (2013).
28. A. M. Auersperg, A. Kacelnik, A. M. von Bayern, Explorative learning and functional inferences on a five-step means-means-end problem in Goffin's cockatoos (*Cacatua goffini*). *PLoS ONE* **8**, e68979 (2013).
29. T. Nakagaki, H. Yamada, Á. Tóth, Maze-solving by an amoeboid organism. *Nature* **407**, 470 (2000).
30. J. H. Reif, "Complexity of the mover's problem and generalizations" in *20th Annual Symposium on Foundations of Computer Science (SFCS 1979)* (IEEE Computer Society, 1979), pp. 421–427.
31. H. Iba, T. Tohge, Y. Inoue, Cooperative transportation by humanoid robots-solving piano movers' problem. *Int. J. Hybrid Intell. Syst.* **1**, 189–201 (2004).
32. T. J. Czaczkes, F. L. Ratnieks, Cooperative transport in ants (Hymenoptera: Formicidae) and elsewhere. *Myrmecol. News* **18**, 1–11 (2013).
33. H. F. McCreery, M. Breed, Cooperative transport in ants: A review of proximate mechanisms. *Insectes Soc.* **61**, 99–110 (2014).
34. D. J. Ozer, Personality, intelligence, and spatial visualization: Correlates of mental rotations test performance. *J. Pers. Soc. Psychol.* **53**, 129 (1987).
35. S. Pietsch, P. Jansen, Different mental rotation performance in students of music, sport and education. *Learn. Individ. Differ.* **22**, 159–163 (2012).
36. E. Fonio *et al.*, A locally-blazed ant trail achieves efficient collective navigation despite limited information. *eLife* **5**, e20185 (2016).
37. A. Gelblum *et al.*, Ant groups optimally amplify the effect of transiently informed individuals. *Nat. Commun.* **6**, 7729 (2015).
38. O. Feinerman, I. Pinkoviezky, A. Gelblum, E. Fonio, N. S. Gov, The physics of cooperative transport in groups of ants. *Nat. Phys.* **14**, 683 (2018).
39. J. E. Ron, I. Pinkoviezky, E. Fonio, O. Feinerman, N. S. Gov, Bi-stability in cooperative transport by ants in the presence of obstacles. *PLoS Comput. Biol.* **14**, e1006068 (2018).
40. O. Ayalon *et al.*, Sequential decision-making in ants and implications to the evidence accumulation decision model. *Front. Appl. Math. Stat.* **7**, 672773 (2021).
41. I. Parberry, *Introduction to Game Physics with Box2D* (CRC Press, 2017).
42. A. Bakar Sayuti Saman, I. Abdramane, Solving a reconfigurable maze using hybrid wall follower algorithm. *Int. J. Comput. Appl.* **82**, 22–26 (2013).
43. M. Nagy *et al.*, Synergistic benefits of group search in rats. *Curr. Biol.* **30**, 4733–4738 (2020).
44. N. Hanusse, D. Ilcinkas, A. Kosowski, N. Nisse, "Locating a target with an agent guided by unreliable local advice: How to beat the random walk when you have a clock?" in *Proceedings of the 29th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing* (Association for Computing Machinery, New York, NY, 2010), pp. 355–364.
45. P. C. Bressloff, Directed intermittent search with stochastic resetting. *J. Phys. A: Math. Theor.* **53**, 105001 (2020).
46. S. Even, *Graph Algorithms* (Cambridge University Press, 2011).
47. K. B. Reed *et al.*, "Haptic cooperation between people, and between people and machines" in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2006), pp. 2109–2114.
48. W. Smoke, R. Zajonc, "On the reliability of group judgments and decisions" in *Mathematical Methods in Small Group Processes* (Stanford University Press Stanford, CA, 1962), pp. 322–333.
49. P. R. Laughlin, *Group Problem Solving* (Princeton University Press, 2011).
50. I. Lorge, H. Solomon, Two models of group behavior in the solution of eureka-type problems. *Psychometrika* **20**, 139–148 (1955).
51. A. Tversky, D. Kahneman, Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science* **185**, 1124–1131 (1974).
52. I. L. Janis, *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes* (Houghton Mifflin, 1972).
53. D. Gigone, R. Hastie, The common knowledge effect: Information sharing and group judgment. *J. Pers. Soc. Psychol.* **65**, 959 (1993).
54. R. H. Sargent, L. S. Newman, Pluralistic ignorance research in psychology: A scoping review of topic and method variation and directions for future research. *Rev. Gen. Psychol.* **25**, 163–184 (2021).
55. J. Y. Halpern, Y. Moses, Knowledge and common knowledge in a distributed environment. *J. ACM* **37**, 549–587 (1990).
56. P. R. Laughlin, A. L. Ellis, Demonstrability and social combination processes on mathematical intellective tasks. *J. Exp. Soc. Psychol.* **22**, 177–189 (1986).
57. A. Newell, *Heuristic Programming: Ill-Structured Problems* (MIT Press, Cambridge, MA, 1993).
58. I. L. Janis, Groupthink. *IEEE Eng. Manage. Rev.* **36**, 36 (2008).
59. R. Baron, N. Kerr, *Group Process, Group Decision, Group Action 2/E* (McGraw-Hill Education, 2003).
60. T. J. Czaczkes, Advanced cognition in ants. *Myrmecol. News* **32**, 51–64 (2022).
61. C. Heins *et al.*, Collective behavior from surprise minimization. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2320239121 (2024).
62. A. Gelblum, I. Pinkoviezky, E. Fonio, N. S. Gov, O. Feinerman, Emergent oscillations assist obstacle negotiation during ant cooperative transport. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14615–14620 (2016).
63. B. López-Mesa, G. Thompson, On the significance of cognitive style and the selection of appropriate design methods. *J. Eng. Des.* **17**, 371–386 (2006).
64. F. Lieder, T. L. Griffiths, Strategy selection as rational metareasoning. *Psychol. Rev.* **124**, 762 (2017).
65. M. Binz, S. J. Gershman, E. Schulz, D. Endres, Heuristics from bounded meta-learned inference. *Psychol. Rev.* **129**, 1042 (2022).
66. W. Toyokawa, Collective cognition and behaviour. *Nat. Hum. Behav.* **7**, 1612–1613 (2023).
67. I. Lorge, H. Solomon, Group and individual performance in problem solving related to previous exposure to problem, level of aspiration, and group size. *Behav. Sci.* **5**, 28–38 (1960).
68. G. Stasser, W. Titus, Pooling of unshared information in group decision making: Biased information sampling during discussion. *J. Pers. Soc. Psychol.* **48**, 1467 (1985).
69. G. Stasser, S. I. Vaughan, D. D. Stewart, Pooling unshared information: The benefits of knowing how access to information is distributed among group members. *Organ. Behav. Hum. Decis. Process.* **82**, 102–116 (2000).
70. A. Shiloni, N. Agmon, G. A. Kaminka, "Of robot ants and elephants" in *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems* (2009), vol. 1. pp. 81–88.
71. M. H. Mabrouk, C. R. McInnes, An emergent wall following behaviour to escape local minima for swarms of agents. *Int. J. Comput. Sci.* **35**, IJCS-35 (2008).