EVOLUTIONARY BIOLOGY

Intra- versus intergroup variance in collective behavior

D. Knebel[1,2], A. Ayali[1,3]*, M. Guershon[1,4], G. Ariel[2]*

Animal collective motion arises from the intricate interactions between the natural variability among individuals, and the homogenizing effect of the group, working to generate synchronization and maintain coherence. Here, these interactions were studied using marching locust nymphs under controlled laboratory settings. A novel experimental approach compared single animals, small groups, and virtual groups composed of randomly shuffled real members. We found that the locust groups developed unique, group-specific behavioral characteristics, reflected in large intergroup and small intragroup variance (compared with the shuffled groups). Behavioral features that differed between single animals and groups, but not between group types, were classified as essential for swarm formation. Comparison with Markov chain models showed that individual tendencies and the interaction network among animals dictate the group characteristics. Deciphering the bidirectional interactions between individual and group properties is essential for understanding the swarm phenomenon and predicting large-scale swarm behaviors.

INTRODUCTION

Cooperative group activity requires a degree of consensus and synchronization. In other words, it is expected that collectivity will result in some homogenization among the individuals forming the group. At the same time, the properties of a coordinated group should somehow be a function of the different traits of the individuals composing it. These general statements bring about ample open questions in biology (1–4), even for simple organisms such as insects in a swarm (5, 6): How do the characteristics of the individual's behavior differ when alone or when in a group? Which traits of the individual are adjusted for it to become part of the synchronized group, which are retained unchanged, and how are they manifested within the swarm? Do the traits of the individual support or interfere with collectivity?

In response to these questions, much research has been devoted to understanding the effect of variability among individuals on the group's collective behavior, both experimentally—ranging from bacteria to primates (7–16)—and theoretically (17–23). See (24–26) for recent reviews and (27, 28) for investigation of heterogeneity in the context of swarm robotics. Part of the interindividual variability has been explained in terms of animal personality—the consistent or context-independent variations in animal behavior [e.g., (29–31)]. Recently, it has been suggested that the inherent differences among members of the group can translate into distinctive group characters (9, 31). Namely, different groups composed of individuals with distinctive features may adopt different collective behaviors. However, the interactions between variability in specific aspects of the individuals' behavior and group-level processes are complex and, moreover, bidirectional, where each level affects and amplifies the other. This leads to a practical difficulty in distinguishing between the inherent variability of the individuals' features and the results of their interaction with the crowd. Accordingly, one of the main goals of the present study was to develop a general methodology for addressing these issues and its application to experiments.

Locusts offer a quintessential example of animal coordinated collective behavior and are therefore exceptionally suited for study of the above questions: Swarms of marching locusts can comprise millions of individuals, aligning or synchronizing their movement across hundreds of square kilometers. Moreover, marching locusts will also demonstrate their distinctive collective behavior under controlled laboratory conditions (4, 6, 32–34) that can partially be reproduced in computer simulations. Although much studied, our knowledge of the complex dynamics and the mechanisms underlying the different aspects of locust collective behavior is far from complete [e.g., (6, 35)]. Moreover, locusts constitute a major threat to human agriculture, which adds to this model organism's particular practical importance.

Here, we used marching locusts under controlled laboratory conditions to study the interdependency between the behavior of individuals and that of the group. To this end, we studied small groups of hoppers, individually tagged with special barcodes, enabling their consistent identification and tracking. The statistics of the behavior of the groups were compared with those of individual locusts introduced singly into the experimental arena and with those of noninteracting, shuffled swarms. The latter were generated by superimposing the trajectories of computer-shuffled real locusts from the experimental groups. These comparisons indicated those behavioral aspects of the individuals that are conserved among groups and necessary for the formation of collective motion. Other individual features, on the other hand, undergo a homogenizing effect by the group but significantly differ between groups, thus generating distinct group characteristics.

Last, using computer simulations, we established that individual differences in social behavioral tendencies can explain the observed variability among groups. In other words, the observed individual heterogeneity leads to the empirical intergroup variance. We applied a simple Markov chain model, which demonstrated that our findings can indeed be explained as resulting from groups composed of unique combinations of locusts that are differing only in their social-behavioral tendencies.

RESULTS

Locusts were individually tagged with barcodes, introduced into a ring-shaped arena (Fig. 1A), either singly ($n = 20$) or in groups of 10 ($n = 20$), and monitored by a video camera for 110 min. After retrieving the position coordinates of each locust in each frame throughout the experiments, a range of kinematic movement parameters and collectivity measures was computed for each group (see Materials and methods for details). Additionally, 20 fictive, shuffled groups were created from the experimental groups by shuffling the locusts' trajectories across the experiments, such that each locust appeared in one shuffled group only (Fig. 1B). This enabled the same

[1]School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel. [2]Department of Mathematics, Bar Ilan University, Ramat-Gan, Israel. [3]Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel. [4]The Steinhardt Museum of Natural History, Tel Aviv University, Tel Aviv 69778, Israel.
*Corresponding author. Email: ayali@post.tau.ac.il (A.A.); arielg@math.biu.ac.il (G.A.)
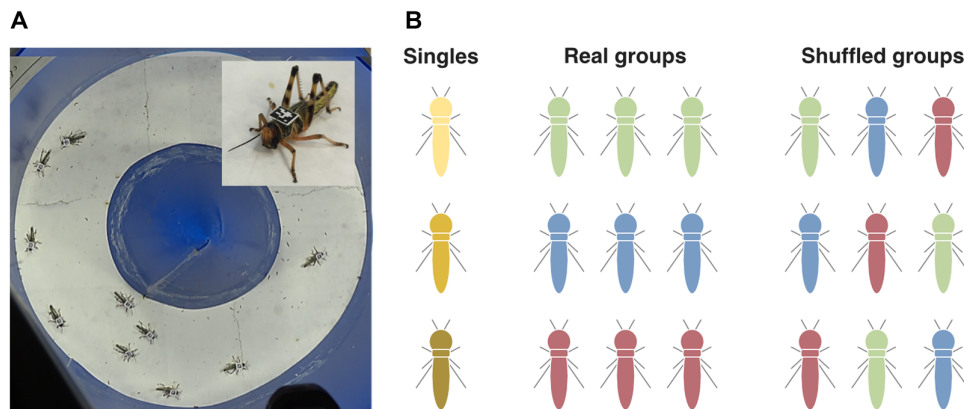
Fig. 1. Experimental procedure. (**A**) A single locust or a group of 10 locusts was introduced into a ring-shaped arena, with a barcode tag attached to each animal's pronotum. Video monitoring and offline analysis enabled following the position of all locusts accurately and consistently throughout the experiment. Photo credit: Daniel Knebel, Tel Aviv University. (**B**) Data comprised three types: single animals in the arena (singles), groups of 10 animals in the arena (real groups), and fictive groups constructed by shuffling the data of the real groups (shuffled groups).

parameters and measures to be calculated also for the shuffled groups, serving as control.

Quantifying collective motion

We first wanted to confirm that the small groups of 10 locusts were moving collectively under our experimental conditions. Collective motion is commonly quantified by measuring the order parameter (36). Here, we define the order parameter as the average direction of moving animals, where the direction is taken as +1 for counterclockwise (CCW) movement, −1 for the clockwise (CW) movement, and 0 for standing. Averaging over all frames in a single experiment (after taking the absolute value), the order parameter could vary between 0 (no preferred direction) and 1 (all moving animals advance in the same direction).

To be precise, denote by $w_i(t)$ the direction in which animal i moves at time t

$$\phi(t) = \frac{1}{Nf(t)} \sum_{i=1}^{N} w_i(t), \ f(t) = \frac{1}{N} \sum_{i=1}^{N} |w_i(t)|$$

where N is the number of walking animals.

As expected for coordinated groups, the order parameter of the real groups was significantly higher than that of the shuffled groups (Fig. 2A; Wilcoxon signed-rank test, $P < 0.001$).

The order parameter is insufficient for differentiating true collectivity from that of a common response of independent individuals to an external stimulus [see also (37)]. To this end, we introduced a new parameter of collectivity, which calculates the mean (over all experiments) of the variance in individual directions in each frame, scaled by the variance expected for independent animals

$$C = \left| 1 - \frac{1}{4p(1-p)} \frac{1}{T} \sum_{t=1}^{T} V(t) \right|$$

where $V(t)$ is the variance in $w_i(t)$ at frame t (among moving animals) and p is the empirical probability to walk in a CCW direction throughout the experiments (0.46). Therefore, $4p(1 - p)$ is the variance of a Bernulli random variable with mean p. The average scaled variance in the direction of walking animals is subtracted from 1, and the absolute value is taken. Thus, we obtain a new scalar parameter, termed the
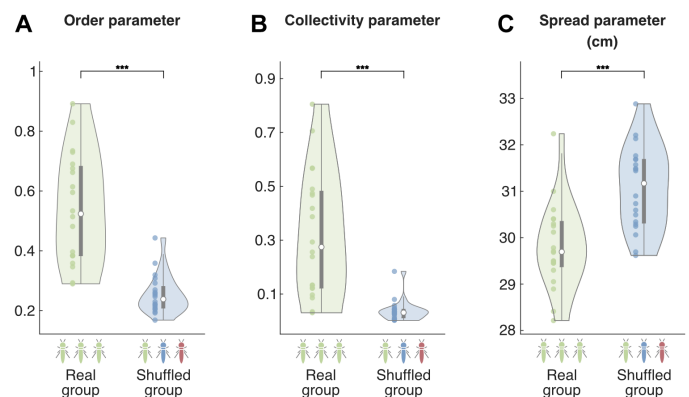

Fig. 2. Collective behavior in groups of 10 locusts. Comparison of collectivity measures between the real and the shuffled groups throughout the experiments. (**A**) Average order parameter. (**B**) Collectivity parameter. (**C**) Spread measure. Each point represents data from a single experiment ($n = 20$). ***$P < 0.001$.

collectivity parameter, which varies from 0 (independent animals) to 1 (a collective swarm). This measure is invariant with respect to a possible bias in the CW/CCW directions (in our experiments, the probability of walking in a CCW direction was found to be $p = 0.46$). Our real groups were found to have a significantly higher collectivity parameter than the shuffled ones (Fig. 2B; Wilcoxon signed-rank test, $P < 0.001$). Another benefit of estimating the collectivity parameter is its insensitivity to fluctuations, which may cause some low, temporary order that is not due to interactions between animals. This is perhaps best emphasized by the almost perfect correlation shown by the collectivity and order parameters in the real groups ($r = 0.91$, Spearman's correlation, $P < 0.001$), compared with the non-significant correlation in the shuffled groups (fig. S1A). Hence, the collectivity parameter is instrumental in distinguishing between the real animal-animal interactions and statistical fluctuations.

Finally, we measured the spatial distribution or the average distance among all animals in each frame, termed the spread measure. The spread measure (for standing and walking insects together) was significantly smaller for the real groups than for the shuffled ones (Fig. 2C; Wilcoxon signed-rank test, $P < 0.001$ for each), indicating the insects' tendency to aggregate (see fig. S2 for results separating standing and walking animals).

Together, the three parameters tested (order, collectivity, and spread; Fig. 2) confirm that 10 locusts in the arena are sufficient for the formation of true collective behavior, which is manifested in collective and ordered marching, as well as in a tendency to aggregate.

The effect of the group on the individual

After establishing that our groups indeed moved collectively, we sought to understand in what respect the social context influences the dynamics of the individuals. We began with relatively simple kinematic measures and compared the small experimental groups with single animals in the same arena.

The fraction of time spent walking, the speed while walking, and the average duration of walking bouts did not significantly differ between the single and the grouped animals (Fig. 3, A to C, respectively). The average pause duration, however, was significantly shorter in the grouped animals (Fig. 3D; Wilcoxon signed-rank test, $P < 0.05$). No significant differences were found in the variances of these parameters between the two experimental conditions: i.e., while in most aspects locust walking was similar whether alone or in a group, among the examined parameters, pause duration was the only feature found to be critical to the formation of collective motion (this does not exclude possible effects of interactions among other measured or unmeasured parameters).

Individual differences within the group

We were interested to learn whether individuals in groups adapt their behavior to others (i.e., retaining the overall statistics of their kinematics). Figure 4A presents the average kinematic measurements in each experiment. As expected, there was no difference in the means of the real and shuffled groups, because, overall, they comprised the exact same individuals and hence the same kinematic measurements, only shuffled. Nonetheless, the variance within the real groups (intragroup variance) was significantly smaller than that within the shuffled ones: the interquartile range (IQR) measured within the real groups across the animals was significantly smaller than that of the shuffled groups, in all kinematic parameters examined (Fig. 4B; fraction of time spent walking, walking speed, walking bout duration, and pause duration; Wilcoxon signed-rank test, $P < 0.01$, $P < 0.001$, $P < 0.05$, and $P < 0.05$, respectively). This indicates that in real interacting groups, there is a homogenizing effect on the group members.

Intergroup variance

The demonstrated homogenizing effect of the group does not necessarily dictate that different groups behave similarly. To examine this, we compared the variance in the kinematics of the real and shuffled groups: If all the groups of locusts were similar, then the differences should not be significant. We found that the variance in the fraction of time spent walking, the walking speed, and the average walking duration was significantly greater for the real groups compared with the shuffled ones (intergroup variance; Fig. 4, Aa to Ac; Brown-Forsythe test, $P < 0.05$, $P < 0.01$, and $P < 0.05$, respectively). Namely, differences between groups were averaged out in the shuffled groups. These findings thus indicate that each real group adopts its own unique characteristics. The average pause duration was again an exception, as there was no significant difference between the variance of the real and the shuffled groups (Fig. 4B). Therefore, while some kinematic quantities can endow each group with unique, distinguishable characteristics, others (i.e., pause duration and the collectivity parameters) were found to be consistent among groups.

The effect of the individual's traits on the group character

Since each group of locusts adopted a unique character, the individual differences among its members, even after the group homogenizing effect (Fig. 4, Ba to Bd), should somewhat determine its nature. To demonstrate a possible interdependency of the group's unique character and the traits of the individuals composing it, we sought to quantify an individual feature or tendency related to the social context that is consistent throughout each experiment.

On the basis of our previous work (4, 34), we hypothesized that the number of individuals walking in the arena is a key stimulus, promoting marching. Accordingly, we calculated the conditional probability of each locust to walk as a function of the number of other walking locusts in the arena (see the example in Fig. 5A). We found that the probability of a locust to walk when five or more other locusts were walking was rather consistent (Fig. 5B), with a correlation of 0.79 between the two halves of the experiment (Spearman's correlation, $P < 0.001$; compare with the non-significant correlation obtained with less than five other walkers, as shown in fig. S3C). This probability (averaged for each experiment) was found to have a high correlation with both the order and the collectivity parameters, reaching 0.71 and 0.7, respectively (Fig. 5C;
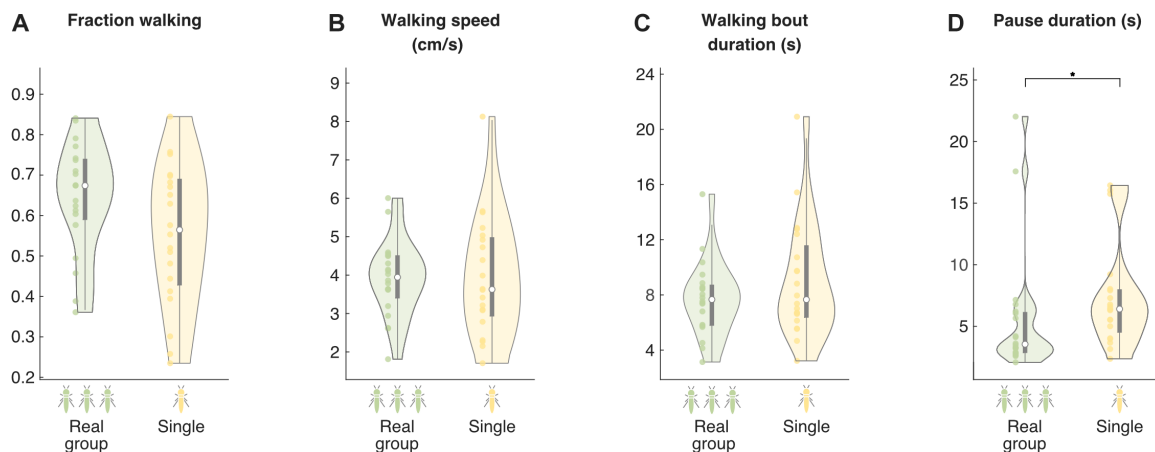

Fig. 3. Kinematic parameters of single locusts versus locusts in real groups. (**A**) The average fraction of time spent walking, (**B**) the average walking speed, (**C**) the average duration of a walking bout, and (**D**) the average pause duration of locusts in real groups and of single locusts. Significant differences are noted between the averages, but not the width of the distributions. *$P < 0.05$.
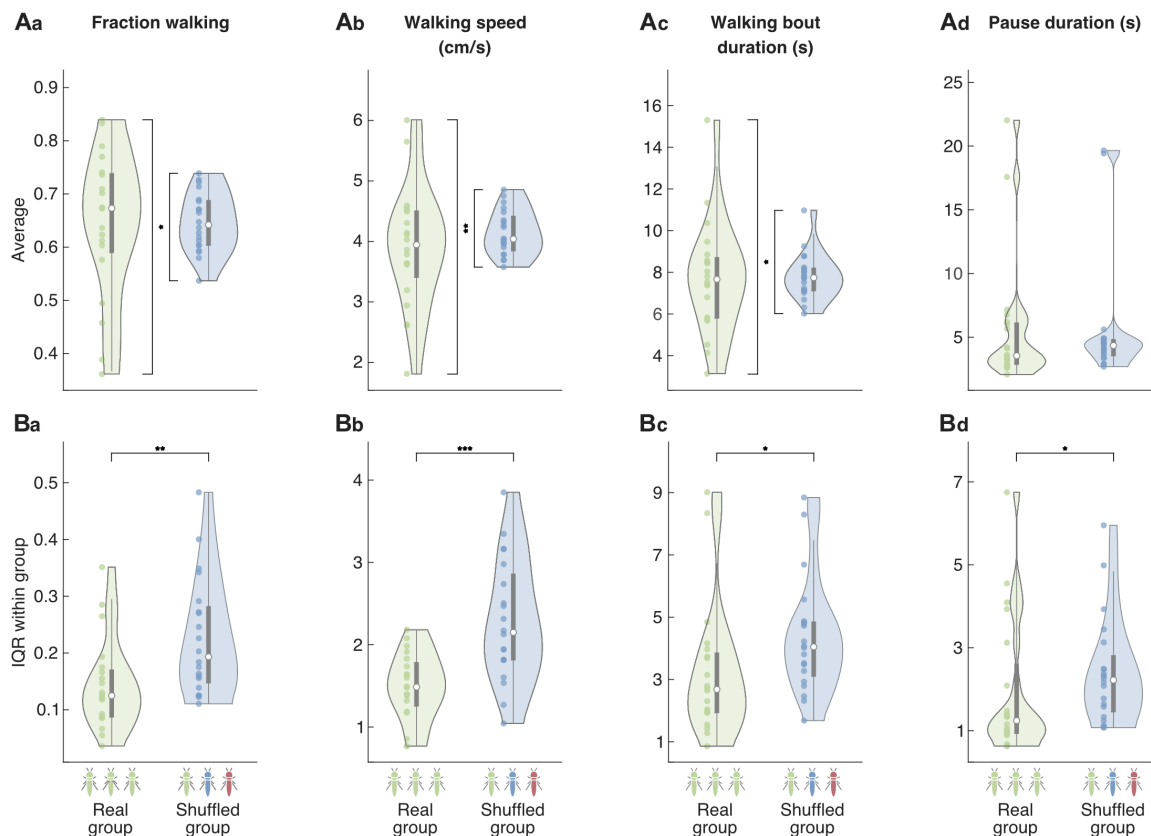
Fig. 4. Kinematics of real versus shuffled groups of locusts. (**Aa** to **Ad**) Comparisons between the average fraction of walking, average walking speed, average walking bout, and average pause duration of the real and shuffled groups, respectively. (**Ba** to **Bd**) Comparisons between the within groups' IQR of the fraction of walking, average walking speed, average walking bout, and average pause duration of the real and shuffled groups, respectively. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.
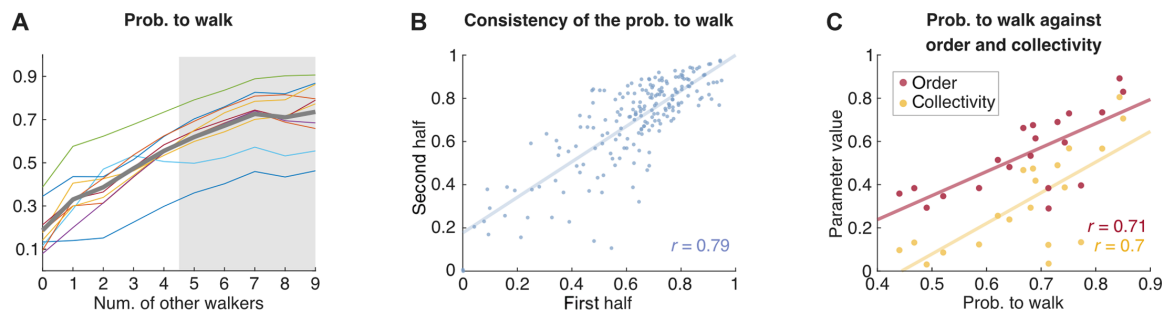

Fig. 5. Kinematics of real versus shuffled groups of locusts. (**A**) Probability of walking as a function of the number of other walkers for a single experiment. Each colored line represents the probability of a specific animal to walk conditioned on the number of other walking animals (0 to 9). The thick gray line shows the experimental average. The shaded area marks the probability of the locusts to walk when five or more others walk (termed P2W$_5$). (**B**) The correlation between the average value of P2W$_5$ obtained using only the first and only the second halves of the experiments. The high correlation suggests that this observable is a consistent individual behavioral characteristic. (**C**) Correlation between the order and the collectivity parameters with P2W$_5$.

Spearman's correlation, $P < 0.001$ and $P < 0.001$, respectively). Similar correlations calculated for the shuffled groups were not significant. These correlations thus demonstrate the interdependency between the characteristics of individuals and the group dynamics.

Mathematical model of individual and group behavior

The correlations described above do not, however, indicate whether the observed intra- and intergroup variances are indeed caused by the heterogeneity in the probability of animals to walk, or vice versa. Therefore,

we designed a mathematical model that incorporated only the minimal individual tendencies described (the individual tendencies to walk when five or more other animals are walking) and explored whether it generated the expected inter- and intragroup variances based on our experimental results.

For a swarm of n animals, the model describes how the system evolves from its state at a given time t, described by $w(t) = (w_1(t), \ldots, w_N(t))$ to its state at a later time $t + 1$, given by $w(t + 1)$. The model assumes an effective coarse-grained discrete time scale (6). Assuming

that it is a homogeneous Markov chain, it is described as a transition probability matrix of dimension $3^N \times 3^N$.

At the beginning of every simulation, we randomly drew the traits of each animal. The traits of animal j determine its probability to change its state (standing, moving CW, or moving CCW). In accordance with the experiments, we assumed that if the fraction of conspecifics walking in the arena is less than 0.5, then the probability of walking is the same for all animals, but grows with f (Fig. 5A). If half the animals or more are walking, then the probability of walking is a random variable, p_j, with values drawn from the empirical distribution described in fig. S3B. As in the experiments, simulated shuffled swarms were created by shuffling simulated single-animal trajectories (with no repetitions). See Materials and methods for further details.

The model generated 1000 independent samples of 10 animal groups (simulated groups), in which each individual received at random one of the real individual tendencies to walk when 5 or more others walk as measured experimentally (overall 200 samples). Shuffling the trajectories produced new simulated-shuffled groups. Both the simulated and simulated-shuffled groups generated order and collectivity parameters, as well as fractions of time spent walking values, comparable with those of the experiments. The order and collectivity parameters, and the fraction of walking time intergroup variance of the simulated groups were higher than those of the simulated-shuffled ones (Fig. 6, A to C). The figure also compares results with homogeneous groups, either only within each group (all animals in a group are the same, but groups are different) or across all groups (all animals in all groups have the average value of p_j). Furthermore, the intragroup variance in the fraction of time spent walking was lower in the simulated groups than in the simulated-shuffled ones (Fig. 6D). These results are all consistent with the experimental data (cf. Figs. 2 and 4) and therefore indicate that the different individual tendencies of locusts are sufficient for the generation of a group's unique characteristics.

Modeling larger swarms

While these simulations mimicked the experimental results, the computational model also enabled us to manipulate the initial distribution of individual traits, the swarm size, and the interaction network among conspecifics. To this end, we studied four model versions corresponding to different interaction networks: one global (each animal interacts with all others) and three local (each animal interacts only with a subpopulation, reflecting the real locust visual field of view; these subpopulations were either fixed or dynamic). See Materials and methods for details.

All local versions showed qualitatively similar dynamics and statistics in almost all parameters (figs. S4 to S7). However, a major difference
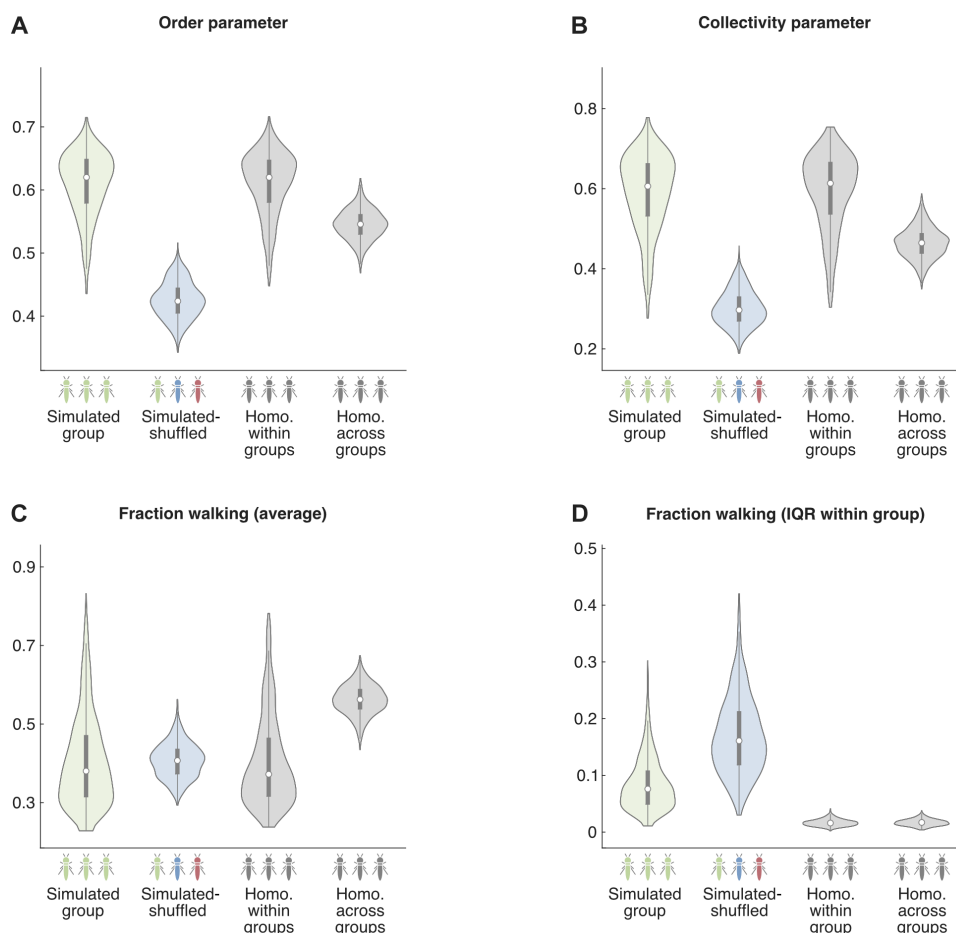

Fig. 6. Simulation kinematic and collectivity measures. The output of computational simulations was compared when the model was introduced with either the data of the locusts in the real experimental groups (real groups), the shuffled data (shuffled groups), and the homogeneous data (same for all group members) equal to the average value of each simulated group (homogenized within groups), or the average of all simulated groups (homogenized across groups). (**A** to **D**) Distribution of the simulated outputs of the order parameter, collectivity parameter, average fraction of walking, and within groups' IQR of fraction of walking, respectively.

was found between the local and global model results. In the latter, the averaged order and collectivity parameters did not decline as the groups grew larger than 100 animals but rather reached an asymptotic value (0.4 to 0.5 for the order parameter and 0.2 to 0.4 for the collectivity). Another result obtained in the local fixed models only was that the intragroup variances of the time spent walking were wider for simulated groups. Yet, they become similar to the shuffled-simulated groups' variance as a function of the group size. The dynamic model, on the other hand, resembled the global model from that perspective.

The combination of results from the various types of models indicates that the intergroup variance is a prominent consequence of the difference among the animals that compose the groups and that it is not masked or averaged out by large groups. Moreover, the intergroup variance carries information on the topology of the social interaction network between conspecifics.

The intragroup variance, however, seems to depend largely on both the kind of information each animal receives and the size of the group. In rather small groups, all model types behave similarly, as the number of animals dictates that all animals receive information about the entire group at each step. However, the differences between the local models indicate that the ability to generate in-group homogeneity is related to the formation of stable subgroups within the swarms. This is demonstrated by the fixed and grid-based models, in which animals receive and deliver information to only a steady subset of the population. There, each subgroup probably generates high uniformity within itself (as estimated for small groups in all models) and reduces the overall swarm inner variability (figs. S4C, S5C, S6C, and S7C).

DISCUSSION

True collective-coordinated behavior is a macroscopic, group-level property, evolving from local interactions among the individual group members [e.g., (1–4)]. Here, we asked whether and how the group affects the behavior of its members and whether individual behavioral tendencies are masked or, on the contrary, manifested in the group behavior.

The overall average and variance of individual locusts' walking kinematics were mostly similar (with the exception of the pause duration), when the locusts were tested singly or as part of a small group. At the same time, the group clearly exerted a homogenizing effect on the kinematics of its members. This is seemingly a contradiction—if the group environment does not affect the overall observed variance, how can variance be reduced by the group? The answer lies in the balance between intragroup homogeneity and intergroup heterogeneity: While each group averages out the properties of the individuals forming it, the average kinematics reached within each group are distinctive. Hence, the specific features of the group are strongly dependent on the combination of its comprising individuals, rather than being determined by the social context per se.

The only notable exception to the above was found in the average duration of pauses between consecutive walking bouts, which was shorter when tested in a group compared with the singles. This behavioral feature, therefore, seems to constitute a fundamental one, most influenced by the animals' social environment. In accord with previous studies that explored the role of intermittent motion in collective behavior [e.g., (4, 38)], the pauses, serving the decision of the individuals to join the collective motion, are critical for the formation of the swarm. The intergroup variance in this particular parameter was rather small and exceptionally similar between the real and the shuffled groups,

again suggesting the importance of the pauses in the collective swarm behavior. We therefore emphasize the importance of examining intergroup variance as an effective tool in identifying those behavioral characteristics that are essential for the generation of collective behavior. Our method is particularly applicable given that it allows studying small swarms, which are typically easier to analyze than large ones.

The consistency of our finding was corroborated by Markov chain simulations. The computational model also enabled us to manipulate the initial distribution of individual traits, the swarm size, and the interaction network between conspecifics. We find that while both the inter- and intragroup variances decrease with swarm size, they also depend on the topology of the graph describing the interaction network within the group.

The above key points regarding collective motion of animal groups in general are even more pertinent for locusts, known for their ability to display density-dependent plasticity in their behavior. Locusts are known to form swarms constituting millions of individuals. This is not to say, however, that our results on small groups of locusts lack natural relevance. The density within a swarm is not fully homogeneous. The collective behavior in low-density areas, such as the outskirts of the swarm, might be greatly determined by the individuals and the local small groups within these areas. Moreover, the coalescence of locusts into destructive plagues commences by means of small-scale local aggregations, followed by a complex and far from fully understood process of phase transformation, further aggregation, and swarming. Similarly, our understanding of how locust swarms disperse is lacking. The data presented here suggest that during both processes (swarm buildup and dispersal), the individual tendencies of the members might play a critical role in the swarm dynamics. These ideas are supported by our simulation results (figs. S4 to S7), for example with the local grid-based model, which indicates that the ability to generate in-group homogeneity is related to the formation of stable subgroups within the swarms.

Finally, our findings highlight the prominence of biological variance, which echoes in every aspect of the life sciences (39, 40). We have demonstrated here a general methodology in which a careful analysis of both individual and group variances can reveal which features are key for the formation of collective motion, and uncover the intricate, synergetic relations between the dynamics of crowds and the personal traits of the individuals compromising it.

MATERIALS AND METHODS
Animals

Desert locusts, *Schistocerca gregaria* (Forskål), were obtained from our colony at the School of Zoology, Tel Aviv University, Israel. The locusts were reared for many consecutive generations under crowded conditions with 100 to 160 individuals in 60-liter aluminum cages under a controlled temperature of 30°C, 35 to 60% humidity, and a 12-hour dark/12-hour light cycle. The locusts were fed daily with wheat seedlings and dry oats. All experiments were performed with nymphs of the final (fifth) nymphal instar (3 to 4 cm in length and ~0.5 cm in width).

Experimental setup

The experimental arena was composed of a flat paper sheet floor circumscribed by an outer blue plastic wall (60-cm diameter by 55-cm height). A circular concentric wall made of similar plastic (diameter, 30 cm) was placed in the center to create a ring-shaped arena (Fig. 1A). The lower 10 cm of the arena walls and central dome were thinly

coated with Fluon (Whitford Plastics Ltd., Runcorn, UK) to prevent the nymphs from climbing. The arena was placed in a room heated to 30°C and lit by three 100-W bulbs. A video camera (Sony FDR-AXP35: 4K Ultra HD) recorded the experiments from above.

Individual locust recognition
Before each experiment, locusts were individually tagged with miniature barcode tags [Fig. 1A; BugTag, Robiotec Ltd., Israel; see also a similar technique using barcodes for tracking insects specially across time in (41)]. Offline analysis of the video recordings by the Robiotec advanced system for consistent and continuous individual identification, complemented by a custom-designed multiple-target tracking and a trajectory-smoothing method [as detailed in (4)], enabled retrieval of the position of each animal throughout the experiment in respect to the arena's center.

The BugTag system enables highly accurate measurements of the tag's center of mass with a resolution of 2 to 3 pixels (corresponding to ca. 0.5 mm) at a rate of 25/3 frames per second. Short segments in which animals were not identified by the system (shorter than 5 cm or 25 s) were interpolated, resulting in about 99% identification. The rest of the frames were analyzed manually, resulting in 100% identification of all animals in the arena.

Experimental conditions
Two types of conditions were tested in the arena: (i) experiments with single locusts ($n = 20$) and (ii) experiments with groups of 10 locusts ($n = 20$). In addition, shuffled groups were created by shuffling the members of the group experiments ($n = 20$). Each locust was used exactly once to create a fictive movie with nine other locusts with which it had not originally swarmed (Fig. 1B).

Analysis of behavior
All data analysis was performed using MATLAB (MathWorks, Natick, MA, USA). The analysis of all experiments was conducted from the 1st to the 111th minute of the recorded movie. Specific attributes of the system and the individual locusts were defined and analyzed as follows:

1) The instantaneous spread measure was defined as the average of all distances between all pairs of animals in each frame. The global spread measure is the average over all frames in a single experiment.

2) The instantaneous walking speed was calculated by the distance an animal traveled over the time of one frame. The global walking speed in an experiment is the average over all animals and frames.

3) Walking bouts and pauses were identified using a repeated running median (RRM) smoothing (4, 42). Walking bouts were defined as segments with RRM speeds greater than 0.25 cm/s for more than 1 s. The global walking bout and pause duration in an experiment are the average duration over all walking bouts or pauses, respectively.

4) Probability of walking as a function of k other walkers was calculated for each locust by the number of frames it walked when k other animals walked divided by the number of frames k other animals walked (Fig. 5A). Denote P2W as the conditional probability to walk, $\text{P2W}_{ik} = P(L_i|O_k)$, where L_i is the event animal i walked and O_k is the event that k other animals walked. To calculate the probability to walk conditioned on the event that K or more (less) others walk, $\text{P2W}_{iK} = P(L_i|M_k)$, the number of frames the animal walked, L_i, and K or more (less) others walk, M_K, was divided by the number of times M_K occurred (Fig. 5B and fig. S3). The global probability to walk when

K or more/less others walked is the average of $\text{P2W}_{iK} = P(L_i|M_k)$ over all animals (Fig. 5C): $\text{P2W}_K = \sum_i \text{P2W}_{iK}/N$.

Statistical analysis
All statistical tests were conducted with MATLAB. To compare between median values, Wilcoxon signed-rank test was used. To compare between variances, Brown-Forsythe test was used. Significant differences in variance were marked using vertical whiskers and asterisks. All correlation values represent Spearman's rank coefficients and are all significant statistically ($P < 0.05$). Violin plots were generated on the basis of the violinplot function for MATLAB provided here: https://github.com/bastibe/Violinplot-Matlab

Markov chain model
We devised a simplified model to test our hypothesis that variability in the response of animals to conspecifics accounts for the observed inter- and intragroup variance in real and shuffled groups. In addition, we applied the model to predict the dynamics of larger swarms.

For a swarm of n animals, the model is essentially a homogeneous Markov chain over the states space $\Omega = \{-1, 0, 1\}^N$. For $w(t) \in \Omega$, we denote $w(t) = (w_1(t), \ldots, w_N(t))$. Recall that $w_i(t) = 0$ implies that animal i is standing at time t, while $w_i(t) = \pm 1$ implies that at time t, animal i is walking in the CCW(+1)/CW(−1) direction. The model assumes a coarse-grained discrete time scale (6) and is determined by the transition probability matrix P of dimension $3^N \times 3^N$. For $u, v \in \Omega$, $P_{u,v}$ is the probability to change from state $u = (u_1, \ldots, u_N)$ to state $v = (v_1, \ldots, v_N)$.

At the beginning of every simulation, we randomly draw the traits of each animal. The traits of animal j determine the probability to have $v_j = -1, 0,$ or 1 as a function of u. In accordance with the experiments, we assume that if the fraction of conspecifics walking in the arena is less than 0.5, then the probability of walking ($v_j = \pm 1$) is $0.1 + 0.6f$. If half the animals or more are walking, then the probability of walking is a random variable, chosen once at the beginning of the simulation for each animal, with values drawn from the empirical distribution described in fig. S3B. If the animal is walking, the probability of walking in the same direction of $\phi(t)$ depends linearly on $\phi(t)$. By fitting to experimental values (4), we take

$$P(\phi(t)v_j > 0) = 0.5 + 0.4|\phi(t)|$$

Simulations were run for 1000 steps with $N = 10, 20, 50, 100, 200, 500, 1000,$ and 2000. The first 100 steps were discarded and not used for the statistics. As in the experiments, simulated shuffled groups were created by shuffling simulated single-animal trajectories (with no repetitions).

Modeling of large swarms
The model described above is global, in the sense that the decision to walk and in which direction depends on all the animals. This assumption makes sense for small swarms, in which the coarse-grained time can be interpreted as the time it takes an animal to "sample" the state of all others. However, within large swarms, a single animal will typically only see a few neighbors and will not have any knowledge of the entire state of the swarm. To this end, we studied three versions of a local model. The model assumes that the probability of an animal to walk depends on how many of the other animals it sees are walking. Let us denote by f the fraction of walking animals it sees, i.e., the number of animals seen walking divided by the number of animals seen (walking or standing). If f is less than 0.5, then the probability to walk in the next

simulation step is $0.1 + 0.6f$. Parameters were obtained by fitting the probability to walk in all 200 tested animals (similar to the left half of Fig. 5A). However, if $f \geq 0.5$, then the probability to start walking is itself random: The values are drawn from the empirical distribution depicted in fig. S3B. They are different for each animal but constant throughout the simulation. The threshold of $f = 0.5$ corresponds to our finding that the influence of five or more walking animals out of nine is indeed an individual trait that is consistent throughout an experiment (Fig. 5B).

The results with the global model, in which every animal sees all others, were detailed above in Results. As in our experimental analysis, simulated-shuffled groups were generated by shuffling animal trajectories between simulation instances. The order and collectivity parameters were computed, as well as the average (among animals in a simulation instance) fraction of time walking. Averages and inter- and intragroup variances (IQR) were compared.

One of the key problems with the global model is the unrealistic assumption that all animals are continuously aware of the walking state of all other animals even within very large groups. As a result, large global swarms fail to synchronize. This can be seen in fig. S4 (Aa and Ab), which shows that both the order and collectivity parameters of simulated swarms become very small as the number of animals grows.

To this end, we studied three versions of local models, in which each animal only sees and reacts to a small number of conspecifics. These versions correspond to different interaction networks among individuals as follows.

1) A fixed local model, in which every animal only sees nine other animals, randomly chosen once at the beginning of each simulation. The set of neighbors is fixed throughout the simulations, i.e., the same nine animals are observed (fig. S5).

2) A dynamic local model, in which every animal only sees nine other animals. A new set of nine neighbors is drawn for each animal every simulation step (fig. S6).

3) A grid-based model, in which animals are initially placed on a square two-dimensional grid. Each animal sees its nearest neighbors on the grid, with periodic boundaries (fig. S7). Note that the number of neighbors in this model is eight.

SUPPLEMENTARY MATERIALS

Fig. S1. Relation between the order and collectivity parameters.
Fig. S2. Aggregation of walking and standing animals.
Fig. S3. Distributions of the social-dependent probability to walk.
Fig. S4. Global model results.
Fig. S5. Local-fixed model results.
Fig. S6. Local-dynamic model results.
Fig. S7. Local grid-based model results.

REFERENCES AND NOTES

1. I. Giardina, Collective behavior in animal groups: Theoretical models and empirical studies. *HFSP J.* **2**, 205–219 (2008).
2. D. J. T. Sumpter, *Collective Animal Behavior* (Princeton Univ. Press, 2010).
3. C. K. Hemelrijk, H. Hildenbrandt, Schools of fish and flocks of birds: Their shape and internal structure by self-organization. *Interface Focus* **2**, 726–737 (2012).
4. G. Ariel, Y. Ophir, S. Levi, E. Ben-Jacob, A. Ayali, Individual pause-and-go motion is instrumental to the formation and maintenance of swarms of marching locust nymphs. *PLOS ONE* **9**, e101636 (2014).
5. I. A. N. Dublon, D. J. T. Sumpter, Flying insect swarms. *Curr. Biol.* **24**, R828–R830 (2014).
6. G. Ariel, A. Ayali, Locust collective motion and its modeling. *PLOS Comput. Biol.* **11**, e1004522 (2015).
7. S. Benisty, E. Ben-Jacob, G. Ariel, A. Be'er, Antibiotic-induced anomalous statistics of collective bacterial swarming. *Phys. Rev. Lett.* **114**, 018105 (2015).
8. I. Fürtbauer, A. Fry, Social conformity in solitary crabs, *Carcinus maenas*, is driven by individual differences in behavioural plasticity. *Anim. Behav.* **135**, 131–137 (2018).
9. I. Planas-Sitjà, J.-L. Deneubourg, C. Gibon, G. Sempo, Group personality during collective decision-making: A multi-level approach. *Proc. R. Soc. B Biol. Sci.* **282**, 20142515 (2015).
10. C. Brown, E. Irving, Individual personality traits influence group exploration in a feral guppy population. *Behav. Ecol.* **25**, 95–101 (2014).
11. J. R. G. Dyer, D. P. Croft, L. J. Morrell, J. Krause, Shoal composition determines foraging success in the guppy. *Behav. Ecol.* **20**, 165–171 (2009).
12. D. R. Farine, A. Strandburg-Peshkin, I. D. Couzin, T. Y. Berger-Wolf, M. C. Crofoot, Individual variation in local interaction rules can explain emergent patterns of spatial organisation in wild baboons. *Proc. R. Soc. B Biol. Sci.* **284**, 20162243 (2017).
13. O. L. Guayasamin, I. D. Couzin, N. Y. Miller, Behavioural plasticity across social contexts is regulated by the directionality of inter-individual differences. *Behav. Processes* **141**, 196–204 (2017).
14. J. E. Herbert-Read, S. Krause, L. J. Morrell, T. M. Schaerf, J. Krause, A. J. W. Ward, The role of individuality in collective group movement. *Proc. R. Soc. B Biol. Sci.* **280**, 20122564 (2013).
15. J. W. Jolles, K. L. Laskowski, N. J. Boogert, A. Manica, Repeatable group differences in the collective behaviour of stickleback shoals across ecological contexts. *Proc. R. Soc. B Biol. Sci.* **285**, 20172629 (2018).
16. J. D. Crall, A. D. Souffrant, D. Akandwanaho, S. D. Hescock, S. E. Callan, W. M. Coronado, M. W. Baldwin, B. L. de Bivort, Social context modulates idiosyncrasy of behaviour in the gregarious cockroach *Blaberus discoidalis*. *Anim. Behav.* **111**, 297–305 (2016).
17. L. M. Aplin, D. R. Farine, R. P. Mann, B. C. Sheldon, Individual-level personality influences social foraging and collective behaviour in wild birds. *Proc. R. Soc. B Biol. Sci.* **281**, 20141016 (2014).
18. G. Ariel, O. Rimer, E. Ben-Jacob, Order–disorder phase transition in heterogeneous populations of self-propelled particles. *J. Stat. Phys.* **158**, 579–588 (2014).
19. K. Copenhagen, D. A. Quint, A. Gopinathan, Self-organized sorting limits behavioral variability in swarms. *Sci. Rep.* **6**, 31808 (2016).
20. L. Guisandez, G. Baglietto, A. Rozenfeld, Heterogeneity promotes first to second order phase transition on flocking systems. arXiv:1711.11531v3 [cond-mat.soft] (30 November 2017).
21. J. W. Jolles, N. J. Boogert, V. H. Sridhar, I. D. Couzin, A. Manica, Consistent individual differences drive collective behavior and group functioning of schooling fish. *Curr. Biol.* **27**, 2862–2868.e7 (2017).
22. A. M. Menzel, Collective motion of binary self-propelled particle mixtures. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **85**, 021912 (2012).
23. S. Mishra, K. Tunstrøm, I. D. Couzin, C. Huepe, Collective dynamics of self-propelled particles with variable speed. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **86**, 011901 (2012).
24. M. del Mar Delgado, M. Miranda, S. J. Alvarez, E. Gurarie, W. F. Fagan, V. Penteriani, Agustina di Virgilio, J. M. Morales, The importance of individual variation in the dynamics of animal collective movements. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, 20170008 (2018).
25. A. P. Modlmeier, C. N. Keiser, C. M. Wright, J. L. L. Lichtenstein, J. N. Pruitt, Integrating animal personality into insect population and community ecology. *Curr. Opin. Insect Sci.* **9**, 77–85 (2015).
26. M. M. Webster, A. J. W. Ward, Personality and social context. *Biol. Rev. Camb. Philos. Soc.* **86**, 759–773 (2011).
27. D. Kengyel, H. Hamann, P. Zahadat, G. Radspieler, F. Wotawa, T. Schmickl, Potential of heterogeneity in collective behaviors: A case study on heterogeneous swarms, in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Springer, 2015), vol. 9387, pp. 201–217.
28. K. Szwaykowska, L. M.-y.-T. Romero, I. B. Schwartz, Collective motions of heterogeneous swarms. *IEEE Trans. Autom. Sci. Eng.* **12**, 810–818 (2015).
29. C. Magnhagen, N. Bunnefeld, Express your personality or go along with the group: What determines the behaviour of shoaling perch? *Proc. R. Soc. B Biol. Sci.* **276**, 3369–3375 (2009).
30. C. N. Keiser, J. N. Pruitt, Personality composition is more important than group size in determining collective foraging behaviour in the wild. *Proc. R. Soc. B Biol. Sci.* **281**, 20141424–20141424 (2014).
31. J. N. Pruitt, L. Grinsted, V. Settepani, Linking levels of personality: Personalities of the 'average' and 'most extreme' group members predict colony-level personality. *Anim. Behav.* **86**, 391–399 (2013).
32. P. E. Ellis, The marching behaviour of hoppers of the African migratory locust (*Locusta migratoria migratorioides* R. & F.) in the laboratory, in *Anti-Locust Bulletin 7* (Anti-Locust Research Centre, 1951).
33. J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, S. J. Simpson, From disorder to order in marching locusts. *Science* **312**, 1402–1406 (2006).
34. G. Amichay, G. Ariel, A. Ayali, The effect of changing topography on the coordinated marching of locust nymphs. *PeerJ* **4**, e2742 (2016).
35. D. A. Cullen, A. J. Cease, A. V. Latchininsky, A. Ayali, K. Berry, J. Buhl, R. De Keyser, B. Foquet, J. C. Hadrich, T. Matheson, S. R. Ott, M. A. Poot-Pech, B. E. Robinson, J. M. Smith, H. Song, G. A. Sword, J. V. Broeck, R. Verdonck, H. Verlinden, S. M. Rogers, From

molecules to management: Mechanisms and consequences of locust phase polyphenism, in *Advances in Insect Physiology* (Elsevier Ltd., ed. 1, 2017), vol. 53.

36. T. Vicsek, A. Zafeiris, Collective motion. *Phys. Rep.* **517**, 71–140 (2012).

37. A. Attanasi, A. Cavagna, L. Del Castello, I. Giardina, S. Melillo, L. Parisi, O. Pohl, B. Rossaro, E. Shen, E. Silvestri, M. Viale, Collective behaviour without collective order in wild swarms of midges. *PLOS Comput. Biol.* **10**, e1003697 (2014).

38. S. Bazazi, F. Bartumeus, J. J. Hale, I. D. Couzin, Intermittent motion in desert locusts: Behavioural complexity in simple environments. *PLOS Comput. Biol.* **8**, e1002498 (2012).

39. S. R. X. Dall, A. M. Bell, D. I. Bolnick, F. L. W. Ratnieks, An evolutionary ecology of individual differences. *Ecol. Lett.* **15**, 1189–1198 (2012).

40. D. L. DeAngelis, *Individual-Based Models and Approaches in Ecology: Populations, Communities and Ecosystems* (CRC Press, 2018).

41. J. D. Crall, N. Gravish, A. M. Mountcastle, S. A. Combes, BEEtag: A low-cost, image-based tracking system for the study of animal behavior and locomotion. *PLOS ONE* **10**, e0136487 (2015).

42. I. Hen, A. Sakov, N. Kafkafi, I. Golani, Y. Benjamini, The dynamics of spatial behavior: How can robust smoothing techniques help? *J. Neurosci. Methods* **133**, 161–172 (2004).