1 **Title:** Emergent behavioral organization in heterogeneous groups of a social insect

2

3 **Authors:** Yuko Ulrich^{a,b,1}, Mari Kawakatsu^{c,1}, Christopher K. Tokita^d, Jonathan Saragosti^a,
4 Vikram Chandra^a, Corina E. Tarnita^{d,2,*}, Daniel J. C. Kronauer^{a,2,*}

5 **Affiliations:**

6 ^aLaboratory of Social Evolution and Behavior, The Rockefeller University, New York, NY
7 10065, USA

8 ^bDepartment of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

9 ^cProgram in Applied and Computational Mathematics, Princeton University, Princeton, NJ
10 08544, USA

11 ^dDepartment of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544,
12 USA

13 ¹Y.U. and M.K. contributed equally to this work.

14 ²C.E.T. and D.J.C.K. contributed equally to this work.

15 ^{*}Daniel J. C. Kronauer, Corina E. Tarnita

16 **Email:** dkronauer@rockefeller.edu (D.J.C.K.), ctarnita@princeton.edu (C.E.T.)

17 Y.U. 0000-0001-9380-7863; M.K. 0000-0003-2072-6653; C.K.T. 0000-0002-8585-4317; V.C.
18 0000-0002-8020-0664; C.E.T. 0000-0003-4476-0012; D.J.C.K. 0000-0002-4103-7729

19

20 **Classification:** Major – Biological Sciences; Minor – Evolution

21 **Keywords:** collective behavior; division of labor; self-organization; response threshold model;
22 clonal raider ant

23 **Author Contributions**

24 YU and DJCK conceived the study. MK, CKT, and CET developed the theoretical approach. YU
25 and DJCK designed the experiments. YU, JS, and VC performed the experiments. YU analyzed
26 the experiments. MK and CKT performed the simulations, and MK, CKT, and CET analyzed the
27 simulation results. MK performed analytical calculations with input from CET. YU, MK, CKT,
28 CET, and DJCK drafted the paper, and all authors provided comments.

29 **This PDF file includes:**

30 Main Text
31 Figures 1 to 5
32 Table 1

Abstract

The composition of social groups has profound effects on their function, from collective decision-making to foraging efficiency. But few social systems afford sufficient control over group composition to precisely quantify its effects on individual and collective behavior. Here we combine experimental and theoretical approaches to study the effect of group composition on individual behavior and division of labor (DOL) in a social insect. Experimentally, we use automated behavioral tracking to monitor 120 colonies of the clonal raider ant, *Ooceraea biroi*, with controlled variation in three key correlates of social insect behavior: genotype, age, and morphology. We find that each of these sources of heterogeneity generates a distinct pattern of behavioral organization, including the amplification or dampening of inherent behavioral differences in colonies with mixed types. Theoretically, we use a well-studied model of DOL to explore potential mechanisms underlying the experimental findings. We find that the simplest implementation of this model, which assumes that heterogeneous individuals differ only in response thresholds, could only partially recapitulate the empirically observed patterns of behavior. However, the full spectrum of observed phenomena was recapitulated by extending the model to incorporate two factors that are biologically meaningful but theoretically rarely considered: variation among workers in task performance efficiency and among larvae in task demand. Our results thus show that different sources of heterogeneity within social groups can generate different, sometimes non-intuitive, behavioral effects, but that relatively simple models can capture these dynamics and thereby begin to elucidate the basic organizational principles of DOL in social insects.

Significance Statement

When individuals interact in an aggregate, many factors that are not known *a priori* affect group dynamics. A social group will therefore show emergent properties that cannot easily be predicted from how its members behave in isolation. This problem is exacerbated in mixed groups, where different individuals have different behavioral tendencies. Here we describe different facets of collective behavioral organization in mixed groups of the clonal raider ant, and show that a simple theoretical model can capture even non-intuitive aspects of the behavioral data. These results begin to reveal the principles underlying emergent behavioral organization in social insects. Importantly, our insights might apply to complex biological systems more generally and be used to help engineer collective behavior in artificial systems.

Main Text

Introduction

The study of collective behavior and self-organization is an active area of research across a diversity of fields, from animal movement (1) to robotics (2), from tissue engineering (3) to public health (4), and from voting (5) to conservation (6). The colonies of social insects in particular are striking examples of highly integrated, complex biological systems that can self-regulate without centralized control (7). Consequently, social insects have emerged as powerful systems to study collective behavior and social dynamics, both experimentally and theoretically (8–12). However, few experimental studies have comprehensively measured the influence of group composition—e.g., in age, genotype, or morphology—on collective organization, because the inherent complexity of many social insect colonies renders their composition intractable. This has limited our understanding of how colony composition affects both individual behavior and emergent group-level organization, and constitutes a major hurdle towards a general and comprehensive systems-level description of social insect colonies.

An emergent colony-level trait that has long been thought to depend on colony composition is division of labor (DOL). DOL describes the non-random variation in task performance among members of a social group (13), and is characterized both by between-individual variation and by individual specialization in task performance. Specifically, DOL has been hypothesized to increase with workforce heterogeneity, based on the observation that individual traits often correlate with individual task allocation (14). For example, workers of different age (15–18), genotype (e.g., patrines (19, 20) or matrines (21)), or morphology (e.g., size (19, 22–24)) can vary in their propensity to engage in tasks such as foraging, nursing, nest construction, or grooming of nestmates.

Experimentally testing this hypothesis in a systematic manner has proved challenging, even as theory has confirmed that workforce heterogeneity can indeed lead to the emergence of DOL (13, 25). One successful theoretical approach relies on the fixed response threshold model (FTM) of task allocation (26, 27). This model assumes that each task has an associated stimulus that signals the colony demand for that task. The magnitude of a given stimulus decreases with the number, efficiency, and/or average time investment of workers performing the corresponding task.

Individuals respond to demands based on internal thresholds that reflect their sensitivity to the

stimulus and govern their likelihood of performing a task given its stimulus level: the higher the stimulus level for a task relative to an individual's threshold, the more likely it is to begin performing the task (see Materials and Methods for a detailed description). Thus, fixed thresholds provide a simple mechanism by which individuals dynamically allocate efforts to meet colony demands.

Previous work on the FTM has focused on differences in individual response thresholds as the primary driver of DOL (26, 28–32). In this simple formulation of the FTM, the heterogeneity in behavior is captured via heterogeneity in individual response thresholds drawn from a normal distribution with mean and variance that can be specific to the task and/or the type of individual. Yet, ants can also vary in other traits, for example in the efficiency with which they perform tasks (33–35) or in the average time spent performing a given task (36). Task demand can be similarly variable: for example, foraging activity levels of workers increase with the number of larvae that they have to tend to (37), and larvae of different genotypes develop into adults with different morphologies when cared for by the same workers (38). Thus, the level of demand emanating from the larvae could depend on their number and genotype. Despite this empirical evidence, few theoretical studies of DOL have explored the significance of inter-individual variation in traits other than response thresholds (14, 39).

Here we combine experimental and theoretical approaches to study the effect of group composition on both individual behavior and colony-level DOL. We use the FTM as a natural starting point, but systematically investigate a suite of parameters that might be associated with different individual traits of interest. To overcome the practical challenges associated with studying complex social systems empirically, we capitalize on the advantages of the clonal raider ant (*Ooceraea biroi*). The unique biology of this species affords unparalleled control over the main aspects of colony composition that are thought to affect individual- and group-level behavior in social insects: genotype, age, and morphology. Specifically, colonies of clonal raider ants are queenless and exclusively composed of workers that reproduce asexually and synchronously, so that all adults within a colony are genetically almost identical and emerge in discrete age cohorts. Furthermore, individuals show variation in ovariole number that is associated with body size and other morphological features (40), making it possible to approximately sort individuals into 'regular workers' (2-3 ovarioles) and 'intercastes' (4-6 ovarioles) based on their size (38). Conveniently, workers of different clonal genotypes, age cohorts, and morphologies can be mixed to create functional chimeric colonies (38). Taking

advantage of these features, we quantify individual and collective behavior of *O. biroi* in response to precise, independent manipulations of colony composition along three independent axes, in a single system, and under standardized conditions.

Results

Baseline theoretical predictions of the ‘simple’ FTM with threshold heterogeneity. To establish baseline predictions in colonies with two types of ants (e.g., of different genotype, age, or morphology), we simulated experimental colonies using the simplest and most commonly-employed formulation of the FTM described above (see also Materials and Methods). Simulated colonies were either pure or mixed with respect to ant type; pure colonies consisted solely of one type of ant or the other, whereas mixed colonies had the two types in equal proportions. The ‘simple’ FTM assumes that the types only differ in mean response threshold. The individual thresholds for each type of ant are drawn from a normal distribution with the corresponding type-specific mean. All other model parameters—task performance efficiency, demand rate, threshold variance—are constant across types. Thus, the only source of heterogeneity in pure colonies was the distribution of individual response thresholds, while in mixed colonies that heterogeneity was compounded by differences in the means of the type-specific distributions. The assumption that some threshold heterogeneity exists even in pure colonies rests on the experimental observation that pure colonies exhibit DOL, yet in the absence of any type of heterogeneity, the FTM cannot produce DOL (32).

In pure colonies, there is a single normal distribution of individual thresholds for a given task. Because individuals from the lower end of the distribution are more sensitive to the stimulus for that task, they tended to perform that task more often than those from the higher end, resulting in DOL. In mixed colonies, there is a bimodal distribution of thresholds for each task, with the thresholds of the two types clustered around the different modes. This wider distribution of thresholds resulted in more pronounced DOL, i.e., both behavioral variation and specialization were greater in mixed colonies compared to pure colonies (Fig. 1a-b).

However, all colonies, irrespective of their composition, had the same mean task performance (Fig. 1c). This is because, while colonies may differ in how they allocate workers to tasks (in this case, within mixed colonies, the two ant types differed in their mean task performance because

the type with the lower average threshold for a given task took up that task more often than the other type), they must perform the same amount of work overall to satisfy a given demand. Thus, on average, colony members spent the same fraction of time performing each task across pure and mixed colonies.

In summary, the simple FTM predicted that (P1) regardless of composition, colonies would exhibit the same average task performance, but that (P2) mixed colonies would exhibit higher overall DOL and that (P3) the two types would behave differently from each other in mixed colonies, but not in their respective pure colonies (see Table 1 for a summary of predictions).

Effects of individual genotype, age, and morphology on individual behavior in experimental colonies. We experimentally tested these theoretical predictions in replicate experimental colonies that were either pure or mixed with respect to genetic, demographic, and morphological composition, manipulating each factor independently from the others (see Materials and Methods, Table S1). For example, demographically pure colonies contained either only young workers (1 month old) or only old workers (3 month old), and mixed colonies contained young and old workers in equal proportions; genotype and morphology were kept constant both within and between these colonies. Similarly, genetically pure colonies contained either only workers of genotype B or of genotype A (see (41) for genotype designations), and mixed colonies contained workers of the two genotypes in equal proportions; age and morphology were kept constant among these colonies. All colonies within an experiment had the same size. Colonies contained 8 or 16 workers—fully functional group sizes in the clonal raider ant—and the same number of age-matched larvae hosted in a Petri dish with a plaster floor (see Materials and Methods). The experiment on genetic effects was performed twice, once with larvae of each genotype.

We used a high-throughput automated tracking system (32) to record and analyze the behavior of all individual ants in 120 experimental colonies. The propensity of each ant to perform extranidal tasks (e.g., foraging, waste disposal) as opposed to intranidal tasks (e.g., nursing) was computed as the two-dimensional root-mean-square deviation (r.m.s.d.) of its spatial coordinates (32) (Fig. 2a; see Materials and Methods). The mean r.m.s.d of a group of ants was used as a proxy for their mean performance of extranidal tasks. To quantify colony-level DOL, we calculated behavioral variation and specialization among colony members. Behavioral variation was computed as the standard deviation across r.m.s.d. values of all ants from the same colony. Specialization was

computed as the mean correlation between individual r.m.s.d. ranks across consecutive days in the experiment (32).

We found that workers of genotype B spent more time away from the nest (i.e., had higher mean r.m.s.d.) than workers of genotype A, both across pure colonies and within mixed colonies (Fig. 2b-c), suggesting a genetic basis for the propensity to perform extranidal tasks (19–21). Old workers spent more time away from the nest than young workers irrespective of colony demographic composition (Fig. 2d). Thus, *O. biroi* displays the classic form of age polyethism typical of social insects (15–18, 42), whereby older individuals allocate more time to extranidal tasks, and younger individuals spend more time at the nest. Finally, regular workers spent less time at the nest than intercastes in mixed colonies, but not across pure colonies (Fig. 2e). Because the larger body size and higher reproductive potential of intercastes correspond to a more queen-like phenotype, these behavioral differences support empirical data from other systems—including other queenless (43) and clonal (44) ant species—where reproductive potential often negatively correlates with foraging activity. Thus, consistent with existing knowledge, our experiments revealed robust differences in behavior (here, the propensity to perform extranidal tasks) across ant genotypes, age cohorts, and morphological types (Fig. 2). Interestingly, however, our experiments showed that different ant types (genotypes and age cohorts, but not morphologies) can have different mean behaviors between the corresponding two types of pure colonies. This is inconsistent with theoretical prediction (P1) that colonies, irrespective of their composition, should have the same mean behavior (Table 1).

Effects of genetic, demographic, and morphological mixing on DOL and individual

behavior in experimental colonies.

We found that, in general, mixed colonies had higher DOL—measured as behavioral variation (Fig. S1) and specialization (Fig. 3)—than pure colonies. Although not all pairwise comparisons were statistically significant, there was no case where pure colonies had significantly higher DOL than mixed colonies. Thus, each of the three forms of workforce heterogeneity tended to promote DOL, consistent with prediction (P2) (see Table 1).

We next assessed the outcome of mixing individuals with different behavioral tendencies on individual behavior. Consider two types of individuals, X and Y. Let X_k and Y_k be the mean behavior of types X and Y, respectively, in pure ($k = p$) or mixed ($k = m$) colonies. We assume that $Y_p > X_p$ and $Y_m > X_m$, to reflect our observation that the type with higher r.m.s.d. in pure

colonies always also had higher r.m.s.d in mixed colonies. Given this assumption, mixing could, in principle, have one of three possible outcomes on individual behavior:

- 1) No effect of mixing on individual behavior: the mean behavioral difference between types across pure colonies is the same as the mean behavioral difference between types within mixed colonies, so that $Y_p - X_p = Y_m - X_m$.
- 2) Behavioral ‘contagion’: individuals of different types become behaviorally more similar on average to each other when mixed, so that $Y_p - X_p > Y_m - X_m$; and
- 3) Behavioral ‘amplification’: individuals of different types become behaviorally more different on average from each other when mixed, so that $Y_p - X_p < Y_m - X_m$.

The simple FTM predicted that the two different types will differ in mean behavior when mixed, but not when in pure colonies (P3) (Fig. 1c), i.e., that behavioral amplification should always be observed. However, in contrast to this theoretical prediction, all three outcomes were observed experimentally: genetic mixing resulted in behavioral contagion (Fig. 2b-c; Student’s *t*-test: $t = 3.86$, $p = 0.002$ in colonies with A brood, $t = 2.62$, $p = 0.02$ in colonies with B brood); demographic mixing had no effect on individual behavior (Fig. 2d; $t = -1.50$, $p = 0.16$); and morphological mixing resulted in behavioral amplification (Fig. 2e; $t = -2.44$, $p = 0.02$).

We further investigated whether mixing had an asymmetric effect on the two ant types, i.e., whether it affected one type more than the other, so that the magnitude of change in type-specific behavior between pure and mixed colonies was different across the two ant types (i.e. $|X_m - X_p| \neq |Y_m - Y_p|$). Testing this hypothesis, we found evidence for asymmetric behavioral contagion in genetically mixed colonies with A brood (Fig. 2b), where mixing affected the behavior of A workers (by increasing their extranidal activity) more than it affected the behavior of B workers (*t*-test $|A_m - A_p|$ vs. $|B_m - B_p|$: $t = 3.86$, $p = 0.0024$). All other scenarios studied displayed symmetric effects of mixing on individual behavior (Fig. 2c: $|A_m - A_p|$ vs. $|B_m - B_p|$: $t = -0.94$, $p = 0.37$; Fig. 2d, $|Young_m - Young_p|$ vs. $|Old_m - Old_p|$: $t = -1.02$, $p = 0.33$, Fig. 2e, $|Regular\ Worker_m - Regular\ Worker_p|$ vs. $|Intercaste_m - Intercaste_p|$: $t = 0.68$, $p = 0.50$).

Thus, both the direction and the magnitude of change in individual behavior between pure and mixed colonies depended on the specific source of workforce heterogeneity.

Theoretical predictions of the extended FTM. The predictions of the simple FTM only partially captured the patterns observed in the experimental colonies (Table 1). Thus, differences in mean

threshold alone were insufficient to explain the observed data, suggesting the need to consider other biologically realistic sources of heterogeneity in the model.

Much like assuming that types differ solely in their threshold means, assuming that types differ only in threshold variance or duration of task performance failed to capture the experimentally observed difference in mean behavior between pure colonies (Fig. S2a-b). However, between-type differences in task performance efficiency alone did reproduce this difference. In fact, if the demand was the same for both tasks, differences in task efficiency were necessary for such a pattern to emerge (SI Appendix).

When types differed only in task performance efficiency, we further found behavioral contagion in mixed colonies, i.e., the types behaved more similarly to each other when mixed. Critically, the asymmetry of this contagion depended on the magnitude of the task demand. If the task demand was not too high, so that both types could keep up with the demand in their pure colonies, then the contagion was always downward (Fig. 4a; analytical results in SI Appendix), i.e., the mixed colony, on average, behaved more like the more efficient type. If, on the other hand, the task demand was so high that the less efficient type could not keep up with task demand in its pure colony, then the contagion could, for certain parameter combinations, be upward (Fig. 4b), i.e., the mixed colony behaved on average more like the less efficient type. Hence, if in addition to differences in task efficiency we also assumed between-type differences in task demand (to reflect possible differences in the intensity of task demand stemming from larvae of different genotypes), we were able to qualitatively recapitulate the asymmetric behavioral contagion observed in genetically mixed colonies. Holding all else fixed, differences in task efficiency guaranteed behavioral contagion; the magnitude of task demand modulated the asymmetry of this contagion, i.e., whether mixed colonies on average behaved more like the more or less efficient type.

Although the combination of between-type differences in efficiency and demand successfully recapitulated the observed behavioral contagion, it failed to capture the other observed effects of mixing on individual behavior, notably instances where mixing had no effect on behavior, or where it resulted in behavioral amplification. If we instead combined the between-type differences in task efficiency with between-type differences in mean threshold (to reflect possible between-type differences in the intrinsic propensity to perform tasks), we were able to qualitatively recapitulate both the effects of demographic mixing (no effect of mixing on type-

specific behavior; Fig. 4c) and the effects of morphological mixing (behavioral amplification in mixed colonies; Figs. 4d, S3). Whether we recapitulated the former or the latter depended on the magnitude of the difference in mean thresholds: a larger difference caused the types to differentiate their behavior more strongly in mixed colonies, leading to the latter; a smaller difference dampened this effect, leading to the former.

In general, the model robustly produced a spectrum of patterns, from behavioral contagion to amplification, across a large parameter space (Fig. 4e). Thus, incorporating additional, biologically realistic sources of heterogeneity into the model led to predictions that qualitatively mirrored the range of empirically observed behavioral patterns, namely: (P4) pure colonies of different ant types can differ from each other and from the mixed colonies in mean behavior, and (P5) mixing two types of ants can lead to behavioral contagion, amplification, or neither (Table 1). Moreover, these extensions preserved prediction P2 (to the extent observed in the experimental data), that mixed colonies tend to have higher DOL than pure colonies (Figs. S4, S5).

Theoretical predictions for mean task performance in non-1:1 mixes. Despite its simplicity, the fixed threshold framework demonstrated remarkable explanatory power in both pure and mixed colonies. Given this success, we used the extended FTM to further explore expected patterns of task allocation in colonies with different ratios of ant types. We focused on the four parameter combinations in Fig. 4 because our analysis showed that they collectively captured all of the patterns observed in the experiments. For each parameter combination, we investigated how the mean task performance of colonies changed as we varied the ratio of the two ant types.

Simulations predicted a striking range of patterns. For the parameter combination that produced no effect in the mixed colonies with equal proportions of the two ant types ('1:1 mixes'), the model produced an approximately linear relationship between mean task performance and the ratio of ant types (Fig. 5a). In all other cases, the mean task performance depended nonlinearly on the ratio of the types. However, the shape of the nonlinear curve differed among the cases. In the cases corresponding to behavioral contagion in the 1:1 mixes, the relationship followed a convex decreasing function, so long as there were enough individuals of the more efficient type such that the colony could keep up with the demand (Fig. 5b; analytical results in SI Appendix); otherwise the colony performed the tasks at a fixed maximum capacity that depended only on the average task duration (Fig. 5c). In the case corresponding to behavioral amplification, the relationship

followed a concave decreasing function (Fig. 5d). Hence, despite one type being more efficient than the other in all cases considered, replacing an individual of the former type with one of the latter type would lead to qualitatively different outcomes depending on the between-type differences in mean threshold.

Regardless of the case studied, the ratio of the types did not alter the qualitative effect of mixing on individual behavior (behavioral contagion, amplification, or no effect); for example, the case that led to behavioral amplification in 1:1 mixes predicted behavioral amplification for all non-1:1 mixes tested (Fig. 5d).

Discussion

By manipulating social group composition along three different axes, we found that the effects of group heterogeneity on behavioral organization vary qualitatively depending on the specific factor under consideration. When ants of two different genotypes, ages, or morphologies were mixed, the inherent behavioral differences between each pair of types were dampened, unaffected, or amplified, respectively. The fact that various sources of heterogeneity that naturally exist in animal groups can have different, and possibly opposing, effects on collective organization underscores the importance of independently considering and controlling them. In nature, as in many experiments, all the factors studied here (larval and worker genotype, age, and morphology), as well as other effects (e.g., environmental conditions, resource availability) will play out simultaneously and in largely intractable ways. Being able to break this complexity down experimentally to study each effect separately and under standardized conditions is unprecedented and provides new insight into the basic organizing principles of behavior in social groups.

The experimental literature on DOL in social insects has historically attributed most inter-individual variation in behavior to variation in response thresholds. Our combined empirical and theoretical analyses, however, suggest that this is only part of the story. Indeed, we found that the simplest and most common implementation of the FTM, which assumes that individuals vary in response thresholds alone, only explained part of the empirically observed patterns of behavior. However, the full spectrum of observed phenomena could be qualitatively recapitulated by extending the model to incorporate heterogeneity in two additional factors: task performance

efficiency and task demand. Both are empirically documented (33–35, 45, 46) but theoretically rarely considered.

Between-type differences in threshold and in task efficiency alone—two sources of heterogeneity with opposing effects on behavioral organization—were sufficient to recapitulate the core of our empirical results. Between-type differences in threshold led to behavioral amplification, making ant types behaviorally more different when mixed than when separated, as is known from previous theoretical work on the FTM (26). In contrast, between-type differences in task efficiency led to behavioral contagion, making ant types behaviorally more similar when mixed than when separated. Our theoretical analysis suggests that the relative strengths of these two sources of heterogeneity might vary with colony composition. In our experiments, varying colony *morphological* composition produced behavioral patterns that were theoretically recapitulated under a relatively strong effect of between-type differences in threshold and a relatively weak effect of differences in efficiency. In contrast, the behavioral patterns observed under varied *genetic* composition matched the theoretical predictions for the case in which differences in efficiency have a relatively stronger effect. Manipulating *demographic* composition corresponded to an intermediate scenario in which the two opposing forces seemed to balance each other out. While both threshold (47–49) and efficiency (33–35) are known to vary with various individual traits in social insects, their relative contributions to age-, genotype- and morphology-based behavioral variation remain poorly understood and deserve further investigation.

A third source of heterogeneity, task demand, was necessary to recapitulate the asymmetry in behavioral contagion (i.e., whether workers in mixed colonies behaved more like one or the other type of workers in pure colonies). Empirically, whether a colony composed of two ant genotypes behaved more like one genotype or the other depended on the genotype of the larvae reared. Coupled with the theoretical analysis, these results suggest that the differences in brood genotype could be a source of differences in task demand. This points to the brood as an important player in the regulation of task allocation, at least for tasks associated with brood care, such as foraging and nursing (50). That the brood can influence colony-level traits has been shown in several social insects where larvae solicit food from workers via chemical (51–53) or behavioral (54, 55) cues that affect worker physiology (52, 56) and behavior (e.g., foraging (57), feeding (58)). However, the effect of larvae on the allocation of tasks across individual workers remains elusive in many social insects due to the challenges associated with measuring individual behavior in groups and precisely controlling brood demand. By taking advantage of automated tracking and

the unique biology of the clonal raider ant, our study overcomes these challenges and advances our understanding of larval factors that affect task allocation: we suggest that brood demand and its effects on task allocation depend not only on the presence and number of larvae (37, 59), but also on larval genotype. These results also provide insights into previous cross-fostering experiments that revealed that interactions between worker and brood genotypes have non-linear effects on brood development (into intercastes vs. regular workers) (38). Our work suggests that these effects might arise, at least in part, because different larval genotypes signal different levels of demand—and thereby differ in the magnitude of their effect on worker behavior—while different worker genotypes differ in their behavioral response to a given level of larval demand. For example, if different larval genotypes solicit food at different rates and different worker genotypes respond differently (e.g., via foraging thresholds or efficiency) to such differences in demand, the interaction between genetically-based larval demand and worker behavioral responses may result in differences in larval nutrition. Such differences may, in turn, lead to the previously reported shifts in larval development and, therefore, adult phenotype (38).

Overall, these findings demonstrate that, despite its simplicity, the FTM has remarkable versatility in recapitulating a broad range of experimental outcomes, while still operating under biologically plausible assumptions. It is important to note, however, that while the behaviors observed are robust and generic—i.e., the parameters chosen to illustrate the versatility of the FTM are representative of large regions of parameter space—little is known about what parameter values might actually correspond to the different experimental types. Nevertheless, even in the absence of such experimental measurements, the model provides a useful starting point to generate testable predictions for increasingly complex colony compositions in the clonal raider ant and possibly other social insects.

Our findings add to the growing literature on the role of individual heterogeneity in the collective behavior of complex biological (e.g., schools of fish, neurons in a brain, pathogen strains sharing a host, etc.) and artificial (e.g., heterogeneous robot swarms, synthetic microbial communities, etc.) systems. Much like colonies of the clonal raider ant, these systems exhibit patterns that can be interpreted as behavioral convergence (60–64), divergence (65), and non-linear effects of mixing on group-level phenotypes (66–68). In turn, these patterns affect important processes such as collective decision-making (5), the transmission and evolution of disease (69, 70), and the evolution of cooperative behavior (71, 72). While different variants of threshold-based models have been employed to study several of these systems (73–76), we still lack a unified theoretical

framework to understand the consequences of individual differences on collective dynamics (77). Thus, a comparative approach to the study of the basic organizing principles of heterogeneous systems across scales constitutes an important next step towards understanding the behavior of complex biological systems.

Materials and Methods

Experimental design. Four experiments were performed to investigate the effect of genetic composition (2 experiments differing in the brood genotype used), demographic composition (1 experiment), and morphological composition (1 experiment). Each experiment comprised three treatments (2 with pure colonies, 1 with mixed colonies). All colonies within one experiment were monitored in parallel, but the different experiments were performed separately.

Experimental colonies were composed of workers of the desired age, genotype and morphology (Table S1), as well as larvae, housed in airtight Petri dishes 5 cm in diameter (corresponding to about 25 ant body-lengths) with a plaster of Paris floor. To control individual genotype, clonally related workers were sourced from the same stock colony. We used two commonly used genotypes, A and B (32, 38, 78). To control individual age, workers were sourced from a single age cohort from the same stock colony. Owing to the synchronized reproduction of *O. biroi*, all age-matched workers collected this way had eclosed within a day of each other. To control individual morphology, age-matched regular workers and intercastes from the same stock colony were screened based on body size (small or large) and the absence or presence of vestigial eyes, respectively. From the time they were collected (1–3 days after eclosion) until the start of experiments, workers of a given type were kept as a group. All workers were tagged with color marks on the thorax and gaster using oil-paint markers. Experimental colonies contained 16 (genetic composition and demographic composition experiments) or 8 (morphological composition experiment) workers and a matching number of age-matched larvae (4-5 days old). This 1:1 larvae-to-workers ratio corresponds to the estimated ratio found in a typical laboratory stock colony in the brood-care phase. We used 8 (genetic composition and demographic composition experiments) or 16 (morphological composition experiment) replicate colonies were set up for each group composition, for a total of 120 colonies.

The experiments took place in a climate room at 25 °C and 75% relative humidity under constant light (*O. biroi* is blind and its behavior is not affected by light). Every 3 days, we cleaned and watered the plaster, and added one prey item (live pupae of fire ant minor workers) per live larva at a random location within the Petri dish.

Behavioral data acquisition and analyses were performed as in (32). Software for image analysis is available at <https://doi.org/10.5281/zenodo.1211644>.

Behavioral data analyses. *O. biroi* colonies switch between reproductive phases, in which all workers stay in the nest and lay eggs, and brood-care phases, in which workers nurse the larvae in the nest but also leave the nest to forage, explore, or dispose of waste. For each colony, behavioral analyses were restricted to the brood-care phase, which started at the beginning of the experiment and ended when all larvae had either reached the non-feeding pre-pupal stage or died.

The spatial distribution of each ant throughout the brood-care phase was quantified as the two-dimensional root-mean-square deviation:

$$\text{r. m. s. d.} = \sqrt{\frac{\sum_i ((x_i - \bar{x})^2 + (y_i - \bar{y})^2)}{n}}$$

in which x_i and y_i are the coordinates of the focal ant in frame i , and $\bar{x}$ and $\bar{y}$ are the coordinates of the center of mass of the focal ant's overall spatial distribution in the brood-care phase, and n is the number of frames in which the focal ant was detected. The r.m.s.d. is bounded between 0 and r , the radius of the Petri dish. Workers that spend a lot of time at the nest with the brood (e.g., nursing the larvae) and little time performing extranidal tasks (foraging or waste disposal) have low r.m.s.d. values, whereas workers that spend more time away from the brood have higher r.m.s.d. values.

For each colony, mean behavior was computed as the average of individual r.m.s.d. values, and behavioral variation was computed as the standard deviation of individual r.m.s.d. values. Both metrics were then averaged across replicate colonies for each treatment.

To quantify specialization, we use a metric appropriate for use on continuous behavioral data (r.m.s.d.). Specialization was defined for each colony as the Spearman correlation coefficient between individual r.m.s.d. ranks on consecutive days, averaged over the brood-care phase. Mean

rank-correlation coefficients were then compared across treatments. For all behavioral analyses, ants were excluded from the dataset if they were detected in less than 30% of the frames acquired within the considered time frame (brood-care phase or day); for ants that died during the brood-care phase, the considered time frame was the portion of the brood-care phase preceding death.

Statistical analyses. Statistical analyses were performed in R (79). Analyses were performed separately for each of the four experiments. As the experiments were performed at different times using different cohorts of ants, we cannot rule out “batch” effects and therefore avoid any statistical analyses comparing treatments across experiments.

Effects of individual attributes traits on behavior. The effects of colony composition (pure, mixed), individual attributes (genotypes A vs. B, Young vs. Old, or Regular worker vs. Intercaste), and their interaction, on individual behavior (individual r.m.s.d.) were investigated using linear mixed effects (LME, function *lmer* of package *lme4*) models with colony as a random factor. If a significant interaction between colony composition and individual attributes was detected, we used a second LME model with a four-level independent fixed variable combining colony composition and individual attributes (X_p , Y_p , X_m and Y_m , where X_k and Y_k are the mean behavior of ant types X and Y, respectively, in pure ($k=p$) or mixed colonies ($k=m$)), followed by a Tukey posthoc test with Bonferroni-Holm correction (function *glht* of package *multcomp*) for the following planned comparisons: X_p vs. X_m , Y_p vs. Y_m , X_p vs. Y_p , and X_m vs. Y_m . When needed, response variables were transformed to satisfy model assumptions.

Effects of genetic, demographic, and morphological mixing on DOL. The effects of the treatment (a 3-level variable: X, Y, and mixed) on division of labor (behavioral variation, specialization) were investigated using generalized linear models (GLM), followed by Tukey posthoc tests with Bonferroni-Holm correction for all three pairwise comparisons.

Effects of genetic, demographic, and morphological mixing on individual behavior. To assess whether type-specific behavior was affected by colony composition, we compared the difference in mean behavior (mean r.m.s.d.) between types across pure colonies to the difference in mean behavior between the same types within mixed colonies (i.e., $Y_p - X_p$ vs. $Y_m - X_m$, where $Y_p > X_p$ and $Y_m > X_m$), using unpaired t-tests, after verifying assumptions of normality. We further tested whether the amplitude of the effect differed across types by comparing the magnitude of change

in type-specific behavior between pure and mixed colonies across the two ant types (i.e. $|X_m - X_p| \neq |Y_m - Y_p|$) with unpaired t-tests, after verifying assumptions of normality.

Theoretical model. The fixed threshold model (FTM) considers a colony of n individuals, N_X of which are of type X and N_Y are of type Y ($N_X + N_Y = n$). Types X and Y represent any pair of the experimentally manipulated sub-colony compositions (i.e., genotypes A and B, Young and Old, or Regular Workers and Intercastes). Without loss of generality, we assume that individuals $1, \dots, N_X$ are of type X and individuals $N_X + 1, \dots, n$ are of type Y. The colony must perform m tasks; for consistency with the experimental approach, we assume that there are two tasks ($m = 2$). At a given time step, an individual can be either performing one of the m tasks (active) or not performing any (inactive). The *task state* of individual i at time t is given by the binary variable $x_{ij,t}$: if individual i is active and performing task j at time t , then $x_{ij,t} = 1$ and $x_{ij',t} = 0$ for all $j' \neq j$; if individual i is inactive and in its rest state, then $x_{ij,t} = 0$ for all j . The task state of the colony at time t is then given by the n -by- m binary matrix $Q_t = [x_{ij,t}]$.

The FTM further assumes that each task j has an associated *stimulus* $s_{j,t}$. This stimulus signals the group-level demand for task j and changes depending on both the rate at which the demand increases (e.g., the demand for foraging increases due to increased hunger in the colony) and the number of individuals performing the task (e.g., the demand for foraging decreases as workers go out and find food). Mathematically, the change in stimulus $s_{j,t}$ is governed by Eq. (1):

$$s_{j,t+1} = s_{j,t} + \delta_j - \frac{\alpha_j^X n_{j,t}^X + \alpha_j^Y n_{j,t}^Y}{n}, \quad (1)$$

where δ_j is the task-specific demand rate, taken to be constant over time; α_j^X and α_j^Y are the task-specific performance efficiency of type X and Y individuals, respectively; and $n_{j,t}^X = \sum_{i=1}^{N_X} x_{ij,t}$ and $n_{j,t}^Y = \sum_{i=N_X+1}^n x_{ij,t}$ are the numbers of type X and Y individuals performing task j at time t , respectively.

Each individual i is assumed to have an internal *threshold* for task j , θ_{ij} , drawn at time $t = 0$ from a normal distribution with mean μ_j and normalized standard deviation σ_j (i.e., expressed as a fraction of the corresponding mean μ_j). Although thresholds may change over the individuals' lifetime, they are assumed to be fixed over the timescale of the experiments and, consequently, over the simulation runs. We refer to μ_j as the *mean task threshold* and to σ_j as the *threshold variance* for task j ; each can be group- and/or task-specific (i.e., α_j^X , α_j^Y , σ_j^X , σ_j^Y).

At each time step, inactive individuals assess the m task stimuli in a random sequence until they either begin performing a task or have encountered all stimuli without landing on a task. For each encountered stimulus, individual i evaluates whether to perform the task by comparing the stimulus level to its internal threshold. Specifically, given stimulus $s_{j,t}$ and internal threshold θ_{ij} , individual i commits to performing task j with probability

$$P_{ij} = \frac{s_{ij}^{\eta}}{s_{ij}^{\eta} + \theta_{ij}^{\eta}}, \quad (2)$$

where parameter η governs the steepness of this response threshold function. The larger the value of η , the more deterministic the behavior; in the limit $\eta \rightarrow \infty$, the response function becomes a step function ($H[s_{j,t} - \theta_{ij}] = 0$ if $\theta_{ij} > s_{j,t}$, 1 if $\theta_{ij} < s_{j,t}$) where H is the Heaviside function). Active individuals spontaneously quit their task with a constant quit probability τ . Active individuals can neither evaluate stimuli nor switch tasks without first quitting their current task.

Each agent-based simulation lasted $T = 10,000$ time steps. All simulations and the subsequent analyses were conducted in R (79).

Acknowledgments. We thank A. Gal for advice on analyses, and O. Feinerman and M. Liu for contributions to the tracking algorithms. Research reported in this publication was supported by grants from the Faculty Scholars Program of the Howard Hughes Medical Institute, the Pew Biomedical Scholars Program, and the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM127007 (D.J.C.K.); Swiss National Science Foundation Advanced Postdoc Mobility (P300P3-147900) and Ambizione (PZ00P3_168066) fellowships, and a Rockefeller University Women & Science fellowship (Y.U.); Army Research Office Grant W911NF-18-1-0325 (M.K.); National Science Foundation Graduate Research Fellowship no. DGE1656466 (C.K.T.); and a Kravis Fellowship (J.S.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Howard Hughes Medical Institute, or the Pew Biomedical Scholars Program. This is Clonal Raider Ant Project paper #13.

References

1. D. J. T. Sumpter, The principles of collective animal behaviour. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 5–22 (2006).
2. J. Werfel, K. Petersen, R. Nagpal, Designing collective behavior in a termite-inspired robot construction team. *Science* **343**, 754–758 (2014).
3. D. J. Cohen, W. James Nelson, M. M. Maharbiz, Galvanotactic control of collective cell migration in epithelial monolayers. *Nat. Mater.* **13**, 409–417 (2014).
4. C. D. Nadell, *et al.*, Cutting through the complexity of cell collectives. *Proc. R. Soc. B Biol. Sci.* **280**, 20122770 (2013).
5. A. J. Stewart, *et al.*, Information gerrymandering and undemocratic decisions. *Nature* **573**, 117–121 (2019).
6. P. A. H. Westley, A. M. Berdahl, C. J. Torney, D. Biro, Collective movement in ecology: from emerging technologies to conservation and management. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, 20170004 (2018).
7. D. M. Gordon, The organization of work in social insect colonies. *Nature* **380**, 121–124 (1996).
8. T. D. Seeley, *et al.*, Stop signals provide cross inhibition in collective decision-making by honeybee swarms. *Science* **335**, 108–111 (2012).
9. A. Khuong, *et al.*, Stigmergic construction and topochemical information shape ant nest architecture. *Proc. Natl. Acad. Sci.* **113**, 1303–1308 (2016).
10. Z.-Y. Huang, G. E. Robinson, Regulation of honey bee division of labor by colony age demography. *Behav. Ecol. Sociobiol.* **39**, 147–158 (1996).
11. A. Brahma, S. Mandal, R. Gadagkar, Emergence of cooperation and division of labor in the primitively eusocial wasp *Ropalidia marginata*. *Proc. Natl. Acad. Sci.* **115**, 756–761 (2018).
12. E. E. Greenwald, L. Baltiansky, O. Feinerman, Individual crop loads provide local control for collective food intake in ant colonies. *Elife* **7**, e31730 (2018).
13. S. N. Beshers, J. H. Fewell, Models of division of labor in social insects. *Annu. Rev. Entomol.* **46**, 413–440 (2001).
14. R. Jeanson, A. Weidenmüller, Interindividual variability in social insects - proximate causes and ultimate consequences. *Biol. Rev.* **89**, 671–687 (2014).
15. D. Naug, R. Gadagkar, The role of age in temporal polyethism in a primitively eusocial wasp. *Behav. Ecol. Sociobiol.* **42**, 37–47 (1998).
16. T. D. Seeley, Adaptive significance of the age polyethism schedule in honeybee colonies. *Behav. Ecol. Sociobiol.* **11**, 287–293 (1982).
17. F. Tripet, P. Nonacs, Foraging for work and age-based polyethism: the roles of age and previous experience on task choice in ants. *Ethology* **110**, 863–877 (2004).
18. B. Hinze, R. H. Leuthold, Age related polyethism and activity rhythms in the nest of the termite *Macrotermes bellicosus* (Isoptera, Termitidae). *Insectes Soc.* **46**, 392–397 (1999).
19. P.-A. Eyer, J. Freyer, S. Aron, Genetic polyethism in the polyandrous desert ant *Cataglyphis cursor*. *Behav. Ecol.* **24**, 144–151 (2012).
20. P. C. Frumhoff, J. Baker, A genetic component to division of labour within honey bee colonies. *Nature* **333**, 358–361 (1988).
21. R. Blatrix, J.-L. Durand, P. Jaisson, Task allocation depends on matriline in the ponerine ant *Gnamptogenys striatula* Mayr. *J. Insect Behav.* **13**, 553–562 (2000).
22. C. L. Kwapich, G. Valentini, B. Hölldobler, The non-additive effects of body size on nest architecture in a polymorphic ant. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, 20170235 (2018).
23. J. Wetterer, The ecology and evolution of worker size-distribution in leaf-cutting ants (Hymenoptera: Formicidae). *Sociobiology* **34**, 119–144 (1999).

- 648 24. G. B. Blanchard, G. M. Orledge, S. E. Reynolds, N. R. Franks, Division of labour and
649 seasonality in the ant *Leptothorax albipennis*: worker corpulence and its influence on
650 behaviour. *Anim. Behav.* **59**, 723–738 (2000).
- 651 25. G. Robinson, Regulation of division of labor in insect societies. *Annu. Rev. Entomol.* **37**,
652 637–665 (1992).
- 653 26. E. Bonabeau, G. Theraulaz, J. L. Deneubourg, Quantitative study of the fixed threshold
654 model for the regulation of division of labour in insect societies. *Proc. R. Soc. B Biol. Sci.*
655 **263**, 1565–1569 (1996).
- 656 27. E. Bonabeau, G. Theraulaz, J.-L. L. Deneubourg, Fixed response thresholds and the
657 regulation of division of labor in insect societies. *Bull. Math. Biol.* **60**, 753–807 (1998).
- 658 28. M. R. Myerscough, B. P. Oldroyd, Simulation models of the role of genetic variability in
659 social insect task allocation. *Insectes Soc.* **51**, 146–152 (2004).
- 660 29. J. Gautrais, G. Theraulaz, J. L. Deneubourg, C. Anderson, Emergent polyethism as a
661 consequence of increased colony size in insect societies. *J. Theor. Biol.* **215**, 363–373
662 (2002).
- 663 30. G. E. Robinson, R. E. Page, “Genetic basis for division of labor in an insect society” in
664 *The Genetics of Social Evolution*, M. D. Breed, R. E. Page, Eds. (Westview Press, 1989),
665 pp. 61–80.
- 666 31. M. Waibel, D. Floreano, S. Magnenat, L. Keller, Division of labour and colony efficiency
667 in social insects: effects of interactions between genetic architecture, colony kin structure
668 and rate of perturbations. *Proc. R. Soc. B Biol. Sci.* **273**, 1815–1823 (2006).
- 669 32. Y. Ulrich, J. Saragosti, C. K. Tokita, C. E. Tarnita, D. J. C. Kronauer, Fitness benefits and
670 emergent division of labor at the onset of group-living. *Nature* **560**, 635–638 (2018).
- 671 33. E. O. Wilson, Caste and division of labor in leaf-cutter ants (Hymenoptera: Formicidae:
672 *Atta*). *Behav. Ecol. Sociobiol.* **7**, 157–165 (1980).
- 673 34. A. Kay, S. W. Rissing, Division of foraging labor in ants can mediate demands for food
674 and safety. *Behav. Ecol. Sociobiol.* **58**, 165–174 (2005).
- 675 35. A. L. Mertl, J. F. A. Traniello, Behavioral evolution in the major worker subcaste of twig-
676 nesting *Pheidole* (Hymenoptera: Formicidae): does morphological specialization influence
677 task plasticity? *Behav. Ecol. Sociobiol.* **63**, 1411–1426 (2009).
- 678 36. A. Weidenmüller, The control of nest climate in bumblebee (*Bombus terrestris*) colonies:
679 interindividual variability and self reinforcement in fanning response. *Behav. Ecol.* **15**,
680 120–128 (2004).
- 681 37. Y. Ulrich, D. Burns, R. Libbrecht, D. J. C. Kronauer, Ant larvae regulate worker foraging
682 behavior and ovarian activity in a dose-dependent manner. *Behav. Ecol. Sociobiol.* **70**,
683 1011–1018 (2016).
- 684 38. S. Teseo, N. Châline, P. Jaisson, D. J. C. Kronauer, Epistasis between adults and larvae
685 underlies caste fate and fitness in a clonal ant. *Nat. Commun.* **5**, 3363 (2014).
- 686 39. R. Jeanson, J. H. Fewell, R. Gorelick, S. M. Bertram, Emergence of increased division of
687 labor as a function of group size. *Behav. Ecol. Sociobiol.* **62**, 289–298 (2007).
- 688 40. F. Ravary, P. Jaisson, Absence of individual sterility in thelytokous colonies of the ant
689 *Cerapachys biroi* Forel (Formicidae, Cerapachyinae). *Insectes Soc.* **51**, 67–73 (2004).
- 690 41. D. J. C. Kronauer, N. E. Pierce, L. Keller, Asexual reproduction in introduced and native
691 populations of the ant *Cerapachys biroi*. *Mol. Ecol.* **21**, 5221–5235 (2012).
- 692 42. P. H. W. Biedermann, M. Taborsky, Larval helpers and age polyethism in ambrosia
693 beetles. *Proc. Natl. Acad. Sci.* **108**, 17064–17069 (2011).
- 694 43. F. Ito, S. Higashi, A linear dominance hierarchy regulating reproduction and polyethism
695 of the queenless ant *Pachycondyla sublaevis*. *Naturwissenschaften* **78**, 80–82 (1991).
- 696 44. A. Bernadou, *et al.*, Stress and early experience underlie dominance status and division of
697 labour in a clonal insect. *Proc. R. Soc. B Biol. Sci.* **285**, 20181468 (2018).
- 698 45. M. Beekman, J. N. M. Calis, W. J. Boot, Parasitic honeybees get royal treatment. *Nature*

699 **404**, 723 (2000).

700 46. A. Dornhaus, Specialization does not predict individual efficiency in an ant. *PLoS Biol.* **6**,
701 2368–2375 (2008).

702 47. C. Detrain, J. M. Pasteels, Caste differences in behavioral thresholds as a basis for
703 polyethism during food recruitment in the ant, *Pheidole pallidula* (Nyl.) (Hymenoptera:
704 Myrmicinae). *J. Insect Behav.* **4**, 157–176 (1991).

705 48. T. Pankiw, R. E. Page Jr., The effect of genotype, age, sex, and caste on response
706 thresholds to sucrose and foraging behavior of honey bees (*Apis mellifera* L.). *J. Comp.*
707 *Physiol. A* **185**, 207–213 (1999).

708 49. T. Pankiw, R. E. Page Jr, Response thresholds to sucrose predict foraging division of labor
709 in honeybees. *Behav. Ecol. Sociobiol.* **47**, 265–267 (2000).

710 50. D. J. C. Kronauer, R. Libbrecht, Back to the roots: the importance of using simple insect
711 societies to understand the molecular basis of complex social life. *Curr. Opin. Insect Sci.*
712 **28**, 33–39 (2018).

713 51. Y. Le Conte, G. Arnold, J. Trouiller, C. Masson, B. Chappe, Identification of a brood
714 pheromone in honeybees. *Naturwissenschaften* **77**, 334–336 (1990).

715 52. A. Maisonnasse, *et al.*, A scientific note on E- β -ocimene, a new volatile primer
716 pheromone that inhibits worker ovary development in honey bees. *Apidologie* **40**, 562–
717 564 (2009).

718 53. F. Mas, M. Kölliker, Maternal care and offspring begging in social insects: chemical
719 signalling, hormonal regulation and evolution. *Anim. Behav.* **76**, 1121–1131 (2008).

720 54. N. Kaptein, J. Billen, B. Gobin, Larval begging for food enhances reproductive options in
721 the ponerine ant *Gnamptogenys striatula*. *Anim. Behav.* **69**, 293–299 (2005).

722 55. K. Kawatsu, Effect of nutritional condition on larval food requisition behavior in a
723 subterranean termite *Reticulitermes speratus* (Isoptera: Rhinotermitidae). *J. Ethol.* **31**, 17–
724 22 (2013).

725 56. B. Oldroyd, T. Wossler, F. Ratnieks, Regulation of ovary activation in worker honey-bees
726 (*Apis mellifera*): larval signal production and adult response thresholds differ between
727 anarchistic and wild-type bees. *Behav. Ecol. Sociobiol.* **50**, 366–370 (2001).

728 57. T. Pankiw, R. E. Page Jr, M. Kim Fondrk, Brood pheromone stimulates pollen foraging in
729 honey bees (*Apis mellifera*). *Behav. Ecol. Sociobiol.* **44**, 193–198 (1998).

730 58. Y. Le Conte, L. Sreng, S. H. Poitout, Brood pheromone can modulate the feeding behavior
731 of *Apis mellifera* workers (Hymenoptera: Apidae). *J. Econ. Entomol.* **88**, 798–804 (1995).

732 59. F. Ravary, B. Jahyny, P. Jaisson, Brood stimulation controls the phasic reproductive cycle
733 of the parthenogenetic ant *Cerapachys biroi*. *Insectes Soc.* **53**, 20–26 (2006).

734 60. N. A. Christakis, J. H. Fowler, The spread of obesity in a large social network over 32
735 years. *N. Engl. J. Med.* **357**, 370–379 (2007).

736 61. A. Berdahl, C. J. Torney, C. C. Ioannou, J. J. Faria, I. D. Couzin, Emergent sensing of
737 complex environments by mobile animal groups. *Science* **339**, 574–576 (2013).

738 62. E. van de Waal, C. Borgeaud, A. Whiten, Potent social learning and conformity shape a
739 wild primate’s foraging decisions. *Science* **340**, 483 LP – 485 (2013).

740 63. P. Broly, J.-L. Deneubourg, Behavioural contagion explains group cohesion in a social
741 crustacean. *PLOS Comput. Biol.* **11**, 1–18 (2015).

742 64. D. Centola, *How Behavior Spreads: The Science of Complex Contagions* (Princeton
743 University Press, 2018).

744 65. V. Bettenworth, *et al.*, Phenotypic heterogeneity in bacterial quorum sensing systems. *J.*
745 *Mol. Biol.* **431**, 4530–4546 (2019).

746 66. S. Kaushik, B. Katoch, V. Nanjundiah, Social behaviour in genetically heterogeneous
747 groups of *Dictyostelium giganteum*. *Behav. Ecol. Sociobiol.* **59**, 521–530 (2006).

748 67. N. J. Buttery, C. R. L. Thompson, J. B. Wolf, Complex genotype interactions influence
749 social fitness during the developmental phase of the social amoeba *Dictyostelium*

750 *discoideum*. *J. Evol. Biol.* **23**, 1664–1671 (2010).
751 68. S. Pande, G. J. Velicer, Chimeric synergy in natural social groups of a cooperative
752 microbe. *Curr. Biol.* **28**, 262–267 (2018).
753 69. A. S. Bell, J. C. De. Roode, D. Sim, A. F. Read, Within-host competition in genetically
754 diverse malaria infections: parasite virulence and competitive success. *Evolution* **60**,
755 1358–1371 (2006).
756 70. A. F. Read, L. H. Taylor, The ecology of genetically diverse infections. *Science* **292**,
757 1099–1102 (2001).
758 71. J. E. Strassmann, Y. Zhu, D. C. Queller, Altruism and social cheating in the social amoeba
759 *Dictyostelium discoideum*. *Nature* **408**, 965–967 (2000).
760 72. S. P. Diggle, A. S. Griffin, G. S. Campbell, S. A. West, Cooperation and conflict in
761 quorum-sensing bacterial populations. *Nature* **450**, 411–414 (2007).
762 73. J. J. Hopfield, Neural networks and physical systems with emergent collective
763 computational abilities. *Proc. Natl. Acad. Sci.* **79**, 2554–2558 (1982).
764 74. P. S. Dodds, D. J. Watts, A generalized model of social and biological contagion. *J.*
765 *Theor. Biol.* **232**, 587–604 (2005).
766 75. A. J. W. Ward, D. J. T. Sumpter, I. D. Couzin, P. J. B. Hart, J. Krause, Quorum decision-
767 making facilitates information transfer in fish shoals. *Proc. Natl. Acad. Sci.* **105**, 6948–
768 6953 (2008).
769 76. P. Melke, P. Sahlin, A. Levchenko, H. Jönsson, A cell-based model for quorum sensing in
770 heterogeneous bacterial colonies. *PLOS Comput. Biol.* **6**, 1–12 (2010).
771 77. J. W. Jolles, A. J. King, S. S. Killen, The role of individual heterogeneity in collective
772 animal behaviour. *Trends Ecol. Evol.* **35**, 278–291 (2020).
773 78. P. R. Oxley, *et al.*, The genome of the clonal raider ant *Cerapachys biroi*. *Curr. Biol.* **24**,
774 451–458 (2014).
775 79. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation
776 for Statistical Computing, 2019).
777

Figures and Tables

Figure 1. Theoretical predictions of the simple FTM with differences in mean task thresholds only.

One hundred replicates were simulated for each colony composition. Each opaque circle represents an individual replicate colony (colony size 16); each solid circle represents the average value across all replicates for its corresponding colony (or sub-colony) composition. Panels show division of labor (behavioral variation (a), specialization (b)) and task performance frequency (c) as a function of colony composition. Type Y individuals have a higher mean threshold than type X individuals for both tasks ($\mu^X = 10$, $\mu^Y = 20$). All other parameters are identical for both types: $\delta = 0.6$, $\alpha = 2$, $\sigma = 0.1$, $\eta = 7$, $\tau = 0.2$ (see Materials and Methods and Table S2 for parameter definitions).

Figure 2. Mean r.m.s.d. (a proxy for mean extranidal activity) as a function of colony composition.

Opaque circles represent mean behavior of individuals in replicate colonies (or sub-colonies). Open circles represent average values across replicate colonies (or sub-colonies). For mixed colonies, data are shown both as type-specific mean behavior (in type-specific colors) and colony-level mean behavior (in ‘average’ color). Identical colors across panels indicate ants of the same genotype, age, and morphological types. Sample sizes indicate the number of replicate colonies. Straight solid black brackets represent the effects of individual traits on behavior (X_p vs. Y_p and X_m vs. Y_m). Black dotted brackets represent the behavioral differences between types in pure ($Y_p - X_p$) and mixed ($Y_m - X_m$) colonies. Black curly brackets represent the effect of mixing on inter-type behavioral differences ($Y_p - X_p$ vs. $Y_m - X_m$). Grey curly brackets represent the asymmetry of the effect of mixing between types ($|X_p - X_m|$ vs. $|Y_p - Y_m|$). **a**: Spatial distribution of two ants with high (blue; genotype B) and low (red; genotype A) r.m.s.d. from the same colony. Arrows point to the corresponding r.m.s.d. values. **b**: Behavior as a function of colony genetic composition in colonies with A

802 brood. Colony size 16. GLMM post hoc Tukey tests (B_p vs. A_p : $z = 7.75$, $p = 3.64 \times 10^{-14}$; B_m vs. A_m : $z =$
803 4.61 , $p = 8.06 \times 10^{-6}$) **c:** Behavior as a function of colony genetic composition in colonies with B brood.
804 Colony size 16. (B_p vs. A_p : $z = 7.45$, $p = 2.80 \times 10^{-13}$; B_m vs. A_m : $z = 7.68$, $p = 6.57 \times 10^{-14}$) **d:** Behavior as a
805 function of colony demographic composition. Colony size 16 ($Young_p$ vs. Old_p : $z = -6.05$, $p = 4.39 \times 10^{-9}$;
806 $Young_m$ vs. Old_m : $z = -13.31$, $p < 2 \times 10^{-16}$). **e:** Behavior as a function of colony morphological composition.
807 Colony size 8. ($Regular\ Worker_p$ vs. $Intercaste_p$: $z = 2.14$, $p = 0.10$, $Regular\ Worker_m$ vs. $Intercaste_m$: $z =$
808 8.95 , $p < 2 \times 10^{-16}$). n.s.: non-significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Figure 3. Specialization (day-to-day rank correlation in r.m.s.d.) as a function of colony

composition. Opaque circles represent replicate colonies. Open circles represent average values across replicate colonies. Sample sizes indicate the number of replicate colonies. Identical colors across panels indicate ants of the same genotype, age, and morphological types. **a:** Specialization as a function of colony genetic composition in colonies with A brood. Colony size 16. (GLM post hoc Tukey tests; B_p vs. mixed: $z = -2.78$, $p = 0.02$; A_p vs. mixed: $z = 1.25$, $p = 0.26$) **b:** Specialization as a function of colony genetic composition in colonies with B brood. Colony size 16. (B_p vs. mixed: $z = -2.41$, $p = 0.047$; A_p vs. mixed: $z = 0.88$, $p = 0.38$) **c:** Specialization as a function of colony demographic composition. Colony size 16 (Young_p vs. mixed: $z = 3.01$, $p = 0.005$; Old_p vs. mixed: $z = 5.01$, $p = 1.63 \times 10^{-6}$) **d:** Specialization as a function of colony morphological composition. Colony size 8 (Regular Worker_p vs. mixed: $z = -4.35$, $p = 4.05 \times 10^{-5}$, Intercaste_p vs. mixed: $z = 2.73$, $p = 0.013$). n.s.: non-significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Figure 4. Theoretical predictions of the FTM on task performance and their robustness. a-d: Task performance frequency as a function of colony composition. One hundred replicates were simulated for each colony composition. Each opaque circle represents a replicate colony (colony size 16); each solid circle represents the average value across all replicates for its corresponding colony (or sub-colony) composition. Horizontal gray lines represent the average value of the pure colonies (first two columns) in their respective panels. Identical colors across panels indicate ants of the same types; in particular, the parameters for type Y ants are fixed across panels **a-d** ($\mu^Y = 10$, $\alpha^Y = 2$). **a-b:** Differences in task efficiency (α) between types and demand rate (δ) across colonies capture asymmetric behavioral contagion, downward (**a**) and upward (**b**). Larvae are more demanding in **b** ($\delta = 1.3$) than in **a** ($\delta = 0.6$). For a given δ , type X_1 is more efficient than type Y ($\alpha^{X_1} = 5$, $\alpha^Y = 2$). Type Y can keep up with demand when the larvae are less demanding (**a**) but not when they are more demanding (**b**); type X_1 can keep up with the demand in both cases. Parameters: $\sigma = 0.1$, $\mu = 10$, $\eta = 7$, $\tau = 0.2$. **c-d:** Between-type differences in task efficiency (α) and mean threshold (μ) capture both a lack of effects from mixing (**c**) and behavioral amplification (**d**). In **c**, type X_2 is less efficient than type Y ($\alpha^{X_2} = 1.5$, $\alpha^Y = 2$) and has a lower threshold for both tasks ($\mu^{X_2} = 7.5$, $\mu^Y = 10$). In **d**, type X_3 is more efficient than type Y ($\alpha^{X_3} = 3$, $\alpha^Y = 2$) and has a higher threshold for both tasks ($\mu^{X_3} = 15$, $\mu^Y = 10$). Parameters: $\sigma = 0.1$, $\eta = 7$, $\delta = 0.6$, $\tau = 0.2$. **e:** Change in between-type relative task performance between mixed and pure colonies (measured as $(Y_m - X_m) - (Y_p - X_p)$) as a function of type X's

840 task efficiency, α^X , and mean task threshold, μ^X . The letters X_1 , X_2 , and X_3 indicate the parameter settings
841 for type X in **a-d**; the blue letter Y indicates the parameter settings for type Y, which are fixed for **a-d** and
842 all grids in **e**. Shades of blue indicate behavioral amplification ($Y_p - X_p < Y_m - X_m$), and shades of brown
843 indicate behavioral contagion ($Y_p - X_p > Y_m - X_m$); light gray indicates regions in which the behavior is
844 undefined according to our definitions of the behavioral patterns, which exclude biologically unrealistic
845 scenarios (see Results). Fifty replicates were simulated for each parameter combination. Parameters: $\eta = 7$,
846 $\sigma = 0.1$, $\tau = 0.2$.

Figure 5. Predictions of the FTM for non-1:1 mixes. Colonies of size 16 with varying ratios of X and Y individuals were simulated under different conditions of threshold values, task-performance efficiency, and task demand. One hundred replicates were simulated for each colony composition. Each large circle represents the mean for that mix of X and Y individuals, while the neighboring smaller circles represent the means of X and Y individuals, respectively, within that mix. The dashed lines indicate the null hypothesis of linear behavioral effects of mixing types. The boxes highlight the behavioral patterns that characterize the 1:1-mixes, and their labels indicate correspondence with panels in Fig. 4 (a with Fig. 4c, b with Fig. 4a, c with Fig. 4b, and d with Fig. 4d). Parameters for each type (X₁, X₂, X₃, Y) are identical to those of the corresponding type in Fig. 4. **a:** X₂ individuals have a lower mean task threshold and are less efficient than Y individuals. **b:** X₁ individuals are more efficient than Y individuals. **c:** X₁ individuals are more efficient than Y individuals, but task demand is high. **d:** X₃ individuals have a much higher mean task threshold than Y individuals and are more efficient.

Table 1. Summary of theoretical results. Theoretical predictions of the simple FTM (top row) and extended FTM (bottom row) for pure and mixed colonies, as well as the pattern of behavioral change observed between them. Text in italic highlights key differences in model predictions. Colors indicate agreement (light green) or disagreement (light red) with experiments.

		Pure colonies		Mixed colonies		Behavioral change from pure to mixed colonies
		Mean task performance	Division of labor (DOL)	Mean task performance	Division of labor (DOL)	
Type of Fixed Threshold Model (FTM)	Simple FTM (with variation in mean response threshold only)	<i>Identical</i> across colonies of different ant types (P1)	Exhibits DOL	<i>Identical</i> to pure colonies (P1)	Higher DOL than in pure colonies (P2)	Behavioral <i>amplification only</i> (P3)
	Extended FTM (with variation in other biologically relevant params)	<i>Different</i> across colonies of different ant types (P4)	Exhibits DOL	<i>Different</i> from pure colonies (P4)	Higher DOL than in pure colonies (P2)	Behavioral <i>contagion, amplification, or neither</i> (P5)

= experiments **agree** with prediction
= experiments **disagree** with prediction

864