# Mechanisms to divide labour in social microorganisms

Guy Alexander Cooper[*1,2], Jorge Peña[3], Ming Liu[2], and Stuart Andrew West[2].

[1]St. John's College, Oxford, OX1 3JP, United Kingdom

[2]Department of Zoology, University of Oxford, Oxford, OX1 3SZ, United Kingdom

[3]Institute for Advanced Study in Toulouse, University of Toulouse Capitole, Toulouse, 31080 Cedex 6, France

**Abstract**

Division of labour, where cooperating individuals specialise to perform complementary tasks, plays a fundamental role at all levels of biological complexity, from genes to complex animal societies. Different labour-dividing species employ different mechanisms to determine how tasks are allocated between individuals, including coordinated, random, and genetic determination of phenotype (caste). It is not clear if this diversity is adaptive—arising because different mechanisms are favoured in different environments—or is merely the result of non-adaptive, historical artifacts of evolution. We analyse theoretically the relative advantages of the two dominant mechanisms employed for reproductive division of labour in social microorganisms: coordinated and random specialisation. We show that coordinated specialisation is more likely to evolve over random specialisation when: (i) social groups are small; (ii) cooperation has a greater relative impact on fitness (i.e., it is more "essential"); and (iii) there is a low cost to coordination. We also show how a fully coordinated group of cells can evolve step-wise from less precise forms of coordination. Our predictions are in broad qualitative agreement with empirical patterns across microorganism species that

[*] Corresponding author: guy.cooper@zoo.ox.ac.uk

divide labour. More generally, our results show how evolutionary theory can help explain both why between-individual coordination evolves, and why there is mechanistic variation between species.

**Key words:** division of labour, random specialisation, phenotypic noise, cellular differentiation, coordination, signalling, social microbes, adaptation, evolutionary theory

## Introduction

Different species use different mechanisms to divide labour between reproductives and their helpers (Figure 1). In some cases, individuals interact with one another to determine which individual will perform which role, leading to a precise allocation of labour (coordinated specialisation). For example, workers feeding royal jelly to larvae to produce honey bee queens (Figure 1A), and how a lack of nitrogen fixation by neighbours causes cells of the bacteria *Anabaena cylindrica* to differentiate into sterile nitrogen-fixing heterocysts (Figure 1B) [1–3]. In other cases, there is no between-individual interaction, and each individual adopts its phenotype randomly (random specialisation). For example, in *Salmonella enterica* co-infections, random biochemical fluctuations within each cell's cytoplasm determines which cells infect the gut tissue and trigger an inflammatory response that eliminates competitor species (Figure 1D) [4,5]. In yet further cases, phenotype is determined by the individual's genotype (genetic control). For instance, in some ant societies whether individuals develop into queens, major, or minor workers can be determined, in part, by their genes (Figure 1C) [6–9]. Across the tree of life some species employ one mechanism to divide labour whereas in other species

mixed forms exist with a combination of coordinated specialisation, random specialisation, or genetic control[6,8,10–13].

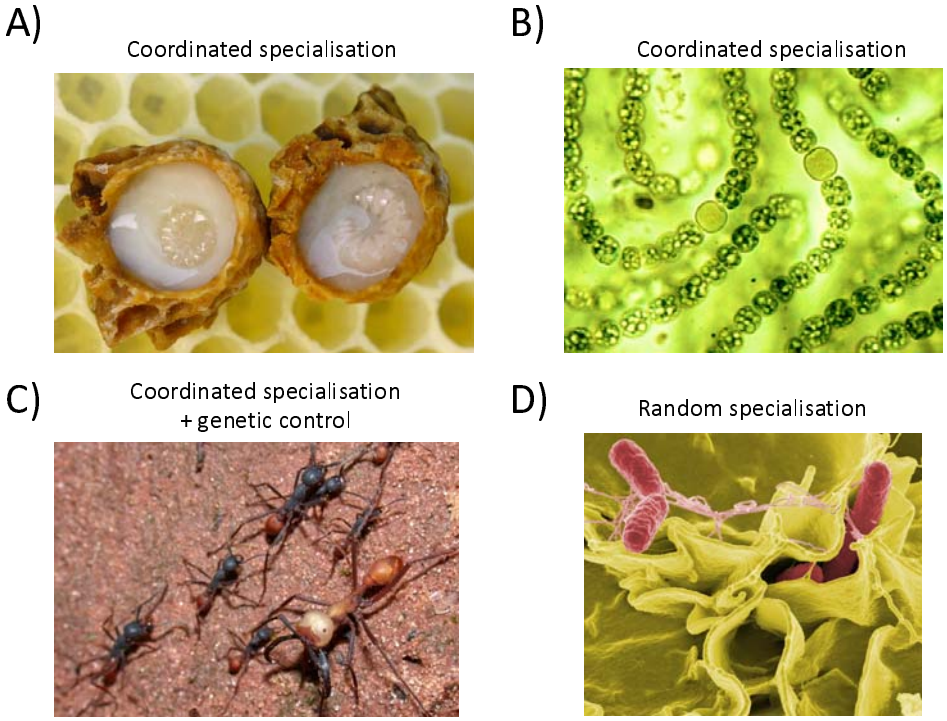


**Figure 1. Different mechanisms for division of labour in nature.** A) In honey bee hives (*Apis mellifera*), larvae develop as sterile workers unless they are fed large amounts of royal jelly by workers (coordinated specialisation) [1]. (Photo by Wausberg via the Wikimedia Commons.) B) In *A. cylindrica* filaments (cyanobacteria), some individuals develop into sterile nitrogen fixers (larger, round cells) if the amount of nitrogen fixed by their neighbours is insufficient (coordinated specialisation). This leads to a precise allocation of labour, with nitrogen-fixing cells distributed at fixed intervals along the filament [2]. (Picture taken by Robert Calentine.) C) In the army ant (*Eciton Burchelli*), whether individual ants become a major or minor worker has a genetic component (genetic control) [9]. (Photo by Alex Wild via the Wikimedia Commons, cropped.) D) In *S. enterica* infections (serovar *Typhymurium*), each cell amplifies intra-cellular noise to determine whether it will self-sacrifice and trigger an inflammatory response that eliminates competing strains (random specialisation). [5] (Photo by Rocky Mountain Laboratories, NIAID, NIH via Wikimedia Commons.)

We lack general evolutionary explanations for why different species use different mechanisms to divide labour between reproductives and their helpers[6,14–16]. Previous

work has focused on non-reproductive division of labour in the social insects, and the proximate mechanisms that lead to different worker castes[9,17–22]. However, the focus in that literature is on a different question – how different proximate mechanisms can produce coordinated specialisation – rather than the broader question of whether coordinated specialisation should be favoured over random specialisation or genetic control in the first place. It is with reproductive division of labour that these three very different mechanisms have been observed in different species[16,23]. There is a lack of theory to explain this variation in the mechanisms used to produce reproductive division of labour[6,14–16].

Bacteria and other social microbes offer excellent opportunities for studying why different labour dividing mechanisms are favoured in different species. Across microbe species, two different mechanisms are used to produce reproductive division of labour - coordinated and random specialisation (Figure 2). Furthermore, while the form of cooperation and life histories of these microbes share many similarities, they also vary in factors that could influence the evolution of division of labour, such as social group size. We develop theory to examine whether the relative advantage of random versus coordinated specialisation can depend upon social or environmental conditions. Our aim here is to use microbes as a ‘test system’ to address the broader question of whether evolutionary theory can explain the diversity in the mechanisms that produce division of labour.

We compare the relative fitness advantage of dividing labour with either coordinated or random specialisation. We begin by assuming that coordinated specialisation leads to social groups containing the optimal allocation of helpers and reproductives. While this assumption is useful as a conceptual overview, such 'full coordination' may require a complex system of mechanisms and could be difficult to evolve from scratch. Consequently, in a secondary analysis we relax the assumption of full coordination and allow the level of coordination to co-evolve along with the extent to which labour is divided. This scenario is more biologically realistic and allows us to examine both whether coordination can initially arise and the extent of coordination favoured.

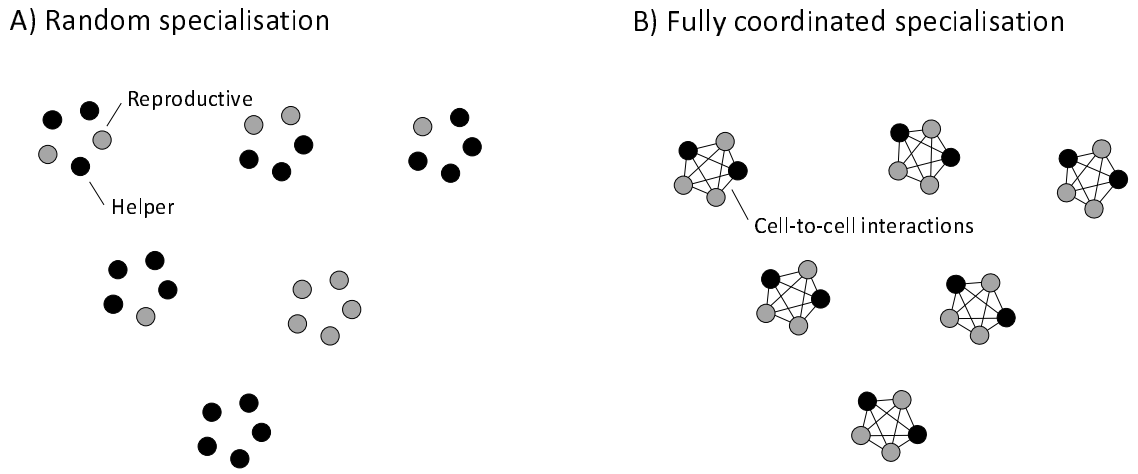


**Figure 2. Mechanisms to divide labour in social microorganisms.** We seek to determine the relative advantages and disadvantages of the two key mechanisms for reproductive division of labour in social microorganisms. (A) Random specialisation occurs when cells randomly specialise into helpers or reproductives independently of one another. This can occur when a genetic feedback circuit is used to amplify small molecular fluctuations in the cytoplasm (phenotypic noise) [4,24–28]. While potentially cheaper as a mechanism, random specialisation can lead to groups that deviate significantly from the optimal proportion of helpers and reproductives. (B) Coordinated specialisation occurs when cells interact with one another, and so share (or gain) phenotypic information while they are differentiating. This can arise through the secretion and detection of extracellular molecules (signals or cues), via the intermediary of

parental control (epigenetics) or through relative condition dependence (signals or cues) [16,29]. While this mechanism can be metabolically costly, it produces precise groups with the optimal proportion of helpers and reproductives.

## Results and Discussion

Our first aim is to capture the problem in a deliberately simple model, which is easy to interpret, and can be applied across diverse microbes[30,31]. We assume a form of cooperation that is common in microbes, when some individuals produce a 'public good' that benefits the other cells. To examine when different mechanisms to divide labour could be favoured, we compare the relative fitness of individuals that use either fully coordinated or random specialization. We then test the robustness of our results, by developing a number of alternate models for different biological scenarios, and by examining how different levels of coordination could evolve.

### Random specialisation vs fully coordinated specialisation

We assume that a single individual arrives on an empty patch and, through a fixed series of replications, spawns a clonal group of $n$ individuals consisting of $k$ sterile helpers and $n-k$ pure reproductives ($k \in \{0,1,2,\ldots,n\}$). The reproductive success (fitness) of a particular group (that we denote by $g_{k,n}$), as measured by the per capita number of dispersing offspring at the end of the group life cycle is given by

$$g_{k,n} = \frac{1}{n}(n-k)f_{k,n}, \qquad (1)$$

where $n-k$ is the number of reproductives in the group, and $f_{k,n}$ is the fecundity of each reproductive, which we assume is increasing in $k$.

Expression (1) highlights the trade-off between the number of reproductives in the group ($n - k$), which is higher when there are fewer helpers (lower $k$), and the amount of help that those reproductives obtain ($f_{k,n}$), which is higher when there are more helpers (higher $k$). The balance of this trade-off often results in an optimal number of helpers, $k^*$, that is intermediate (i.e., $0 < k^* < n$).

In species that divide labour by coordination, we assume that group members share phenotypic information via signalling between cells, parental control, or some cue of relative condition (Figure 2B). Our model is deliberately agnostic to the details of how phenotypic information is shared so as to facilitate broad predictions across different systems. Coordination allows the outcome of individual specialisation to depend on the phenotypes of social group neighbours. In our first analysis, we make the simplifying assumption that individuals coordinate fully, so that groups always form with the optimal number of helpers, $k^*$. The disadvantage of coordinated specialisation is that the mechanism of coordination can incur metabolic costs (e.g., from the production of extracellular signalling molecules). We posit that the fitness of a group of coordinated specialisers is given by:

$$w_C = (1 - c_n) g_{k^*,n} \ , \qquad (2)$$

where $g_{k^*,n}$ is the fecundity of a group with the optimal number of helpers (i.e., $g_{k^*,n} = \max_k g_{k,n}$) and $0 \leq c_n \leq 1$ is a multiplicative cost of coordination that we assume is increasing in group size, $n$. A number of different models have examined how different proximate mechanisms could be favoured to coordinate division of labour in specific systems [17,20–22].

In species that divide labour by random specialisation, each individual in the group independently becomes a helper with a given probability and a reproductive otherwise (Figure 2A). Hence, the final number of helpers in the group is a random binomial variable. Assuming that the probability of adopting a helper role is equal to the optimal proportion of helpers ($p^* = k^*/n$), the expected fitness of a group of random specialisers is given by:

$$w_R = \sum_{k=0}^{n} \binom{n}{k} p^{*k} (1 - p^*)^{n-k} g_{k,n}. \quad (3)$$

The potential advantage of random specialisation is that there are no upfront metabolic costs from, for example, between cell signalling. The downside of random specialisation is that groups sometimes form with fewer or more helpers than is optimal (demographic stochasticity).

In a notable exception to the lack of previous work in this area, Wahl examined when different mechanisms could be favoured to produce non-reproductive division of labour between different castes [15]. Analogous to the cost of random specialization that we have demonstrated, she showed a mechanistically different cost that arises when division of labour is determined genetically: groups may sometimes form that do not contain all of the genotypes required to produce all of the necessary phenotypes (castes) [15].

Before proceeding, we must specify how reproductive fecundity depends on the amount of help in the group. We focus here on one of the most common forms of cooperation in

microbes, where individuals secrete factors that provide a benefit to the local population of cells ("public goods") [32]. We assume that the amount of public good in the social group depends linearly on the number of helpers in the group and is "consumed" by all group members equally [33,34]. An example of such a public good is found in *Bacilus subtilis* populations, where helper cells produce and secrete proteases that degrade proteins into smaller peptides, which are then re-absorbed as a nutrient source by all cells [35].

We consider the possibility that the importance of producing public goods may vary. Accordingly, the fecundity of a reproductive is modelled as the sum of a baseline reproduction $(1-\epsilon)$ and of benefits derived from the production of public goods by helpers (scaled by $\epsilon$). The parameter $\epsilon$, which we call the essentiality of cooperation, quantifies the degree to which reproductives are reliant on helpers in the group [36]. As $\epsilon$ varies from zero to one, the production of public goods goes from being completely unimportant to entirely essential. These assumptions lead to the following expression for the fecundity of a reproductive:

$$f_{k,n} = (1-\epsilon) + \epsilon\frac{k}{n}. \tag{4}$$

Substituting Equation 4 into Equations 1-3 and solving for the condition for fully coordinated specialisation to be favoured over random specialisation (i.e., $w_C > w_R$), gives the simplified expression (see supplementary section C):

$$c_n < \frac{2\epsilon-1}{n}, \tag{5}$$

where division of labour is favoured to evolve only when $\epsilon > 1/2$, which we will assume henceforth. Condition (5) gives three key (and easy to interpret) predictions for how the metabolic cost of coordination ($c_n$), the size of the group ($n$), and the essentiality of cooperation ($\epsilon$) affect which labour-dividing mechanism is optimal (Figure 3).

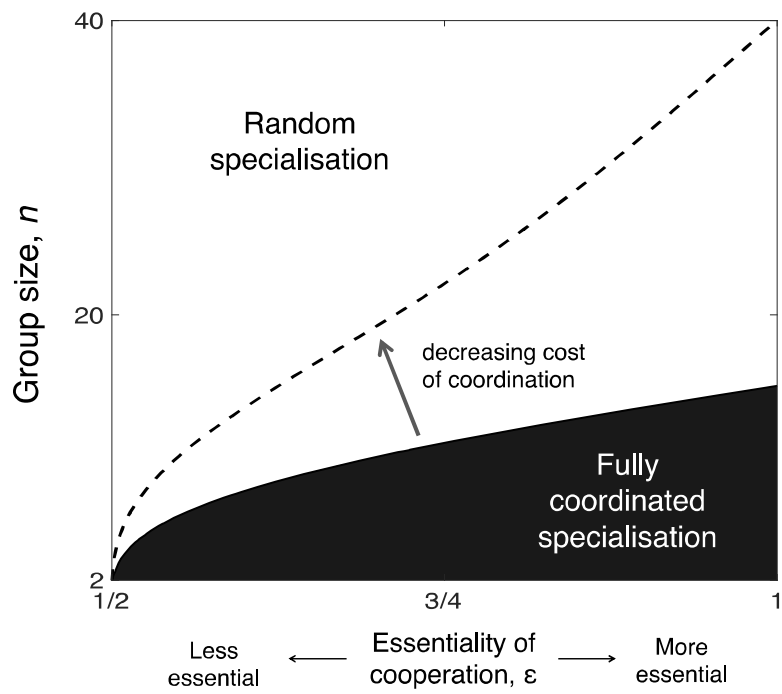


**Figure 3. Random specialisation versus fully coordinated specialisation.** We find that small group sizes (lower $n$), more essential cooperation (higher $\epsilon$), and low costs to coordination (lower $c_n$) favour division of labour by fully coordinated specialisation (dark) over division of labour by random specialisation (white). We modelled the cost of coordination as the increasing and decelerating function $c_n = \alpha\left(1 - e^{-\beta n(n-1)/2}\right)$, where $n(n-1)/2$ is the number of cell-to-cell interactions in a coordinated group and $0 < \alpha < 1$ and $\beta > 0$ are parameters affecting the scale and the shape of the cost of coordination, respectively. Here, $\beta = 0.01$, and $\alpha = 0.1$ (solid boundary) or $\alpha = 0.025$ (dashed boundary).

If division of labour is favoured ($\epsilon > 1/2$) and there is no coordination cost for all $n$ then coordinated specialisation is trivially always the preferred strategy (i.e., $w_C > w_R$ always holds). For positive coordination costs, condition (5) predicts:

*Prediction 1: Smaller costs of coordination favour coordinated specialisation.*

The relative costs of coordination will depend on the specific mechanism by which individuals ensure that the optimal allocation of labour is achieved. For example, signalling between social partners (as occurs in the formation of *Dictyostelium discoideum* fruiting bodies) could have higher costs than parental control (as may occur in Volvocine algae) or than using the products of cooperation as a cue (as occurs in *A. cylindrica)* [37–41].

If there is a cost to coordination, then the optimal mechanism to divide labour depends on how this cost balances against the hidden costs of random specialisation (i.e., the right hand side of inequality (5)). Theses hidden costs are determined by: (i) the likelihood that random groups deviate from the optimal proportion of helpers, and (ii) the degree to which those deviations from the optimal proportion of helpers leads to a reduced fecundity for the group (see supplementary section C). As captured by condition (5), the importance of these two factors depends upon the size of the group (*n*) and the essentiality of cooperation ($\epsilon$).

*Prediction 2. Smaller social groups favour coordinated specialisation.*

Smaller groups favour coordinated specialisation for two reasons. First, the hidden costs of random specialisation are greater in smaller groups. In smaller groups, there are fewer possible outcomes for the proportion of helpers ($n + 1$ possible allocations of labour for groups of size $n$). Consequently, random specialisation can more easily lead to the formation of groups with a realized proportion of helpers that deviates significantly from the optimum ($p \ll p^*$ or $p \gg p^*$). In contrast, in larger groups there are more possible outcomes and the resulting proportion of helpers will be more clustered about the optimal composition with highest fitness (with $p \approx p^*$ for very large group sizes).

This effect of group size on the cost of random specialisation is a consequence of the law of large numbers (e.g., one is much more likely have outcomes close to 50% heads when tossing 100 coins in a row compared to tossing 4 coins in a row, in which no heads or all heads may frequently occur). Our prediction here is also analogous to how, when mating occurs in small groups, small brood sizes select for more precise and less female biased sex ratios to decrease the probability of producing a group containing no males [42–44]. Similarly, in some cases random specialisation in a small group may strategically favour a lower (or higher) proportion of helpers than might otherwise be optimal, to decrease the chances of having zero or insufficient reproductives (or helpers; see supplementary section F and Figure S4).

The second reason why smaller groups favour coordinated specialisation is that in smaller groups there are fewer individuals with whom to coordinate and thus a smaller cost of coordination, given our assumption that $c_n$ is increasing in $n$. The importance of

this influence of group size depends upon mechanistic details about how coordinated is achieved.

*Prediction 3: More essential public goods favour coordinated specialisation*

When cooperation is more essential (larger $\epsilon$), the hidden costs of random specialisation increase, and so coordinated specialisation is more likely to be favoured. As the essentiality of cooperation increases, the fitness costs incurred from producing too few helpers also increases (i.e., there are larger costs to deviating from optimum). In addition, as cooperation becomes more essential (larger $\epsilon$), the fraction of helpers favoured increases towards 50% helpers ($p^* \approx \frac{1}{2}$). This increases the variance in the proportion of helpers produced by random specialisers, and so sub-optimal groups may arise more frequently (see supplementary section C). Both of these effects increase the hidden costs of random specialisation, and thus favour the evolution of fully coordinated specialisation (Figure 3).

<u>Alternative forms of cooperation</u>

The above analysis employs a deliberately simple public goods model, focusing on factors that are expected to be broadly relevant across many microbial systems. This facilitates the interpretation of our results, and generates broadly applicable predictions that are less reliant on the specifics of any particular biological species.

In order to test the robustness of our results (predictions 1-3) we also developed a number of alternate models for different biological scenarios (see supplementary

sections D and E, and Figures S1-S3). In particular, we examined the possibility that the collective good provided by helpers: (i) is not consumed by its beneficiaries, as may occur when self-sacrificing *S. enterica* cells enter the gut to trigger an immune response that eliminates competitors (i.e., a non-rivalrous or non-congestible collective good); or (ii) is only consumed by the reproductives in the group, as may occur for the fixed nitrogen produced by heterocyst cells in *A. cylindrica* filaments (i.e., a kind of excludible good) [4,45,46]. We also examined the possibility that reproductive fecundity depends non-linearly on the proportion of helpers in the group. In all of these alternative biological scenarios, we found qualitative agreement with the predictions of the linear public goods model.

That said, we made a number of assumptions that it would be useful to investigate theoretically. We assumed that social groups are always clonal, such that natural selection would act to maximise group fitness [47,48]. Relaxing this assumption of clonality would both decrease the optimal level of cooperation and increase the potential for conflict within groups, which could undermine the stability and honesty of inter-cellular signalling, making coordination less likely to be favoured in some cases [36,38,49]. Analogously, increasing within-group conflict could favour a genetic predisposition (some genetic control) to adopting a reproductive role in systems that otherwise employ coordinated or random specialisation [6,8].

<u>The optimal level of coordination</u>

Finally, we relax our assumption that individuals coordinate fully, and allow for intermediate forms of coordination (Figure 4A). This allows us to examine whether costly coordination can evolve from an ancestral state where there is no coordination. We assume that coordination is achieved by one-to-one interactions between specialising cells whereby phenotypic information is shared. We let $s$ be the level of coordination, which is equal to the probability that any two cells interact with one another. Once again, our model is agnostic to the specific mode of interaction (whether it be via an extracellular signal, or the intermediary of parental control, or otherwise) so as to generate predictions that are not anchored to the specifics of a particular system. Cells randomly take turns specialising, and adopt a helper role if the proportion of helpers amongst the cells with which they interact falls below a critical threshold, $t$ (see supplementary section G.1 for more details on this model). We assume that higher levels of coordination (higher $s$) entail a larger cost of coordination $c_n$. Overall, this leads to a trade-off between more costly and precise forms of coordination on the one hand (higher $s$) and cheaper and less precise forms of coordination on the other hand (lower $s$). At the two extremes of $s$, we have either random specialisation with no coordination ($s = 0$), or full coordination between all cells ($s = 1$) (Figure 4A).

We used individual based simulations to determine the optimal level of coordination, $s^*$, and target proportion of helpers, $t^*$, that co-evolve in different scenarios (see supplementary section G.1, Figure 4B-D and Figure S5). The evolutionary outcomes, $s^*$ and $t^*$, are estimated by the average trait values that evolved across the entire population, in the last 3,000 generations of 10 independent simulations. The results of

our simulations are in broad qualitative agreement with our previous predictions: small groups (low $n$) and more essential cooperation (high $\epsilon$) favour the evolution of coordinated specialisation ($s^* > 0$; Figure 4D). Additionally, we found that both intermediate forms of coordination ($0 < s^* < 1$) and fully coordinated specialisation ($s^* \approx 1$) can be favoured to evolve, the latter occurring when group size is particularly small and cooperation is particularly essential (lower $n$ and higher $\epsilon$; Figure 4D). We find the same qualitative results with a simple analytical model (see supplementary section G.3). As well as potentially explaining variation across systems, these results also suggest that individual cells could be selected to conditionally adjust their level of coordination in response to factors such as changing group size.

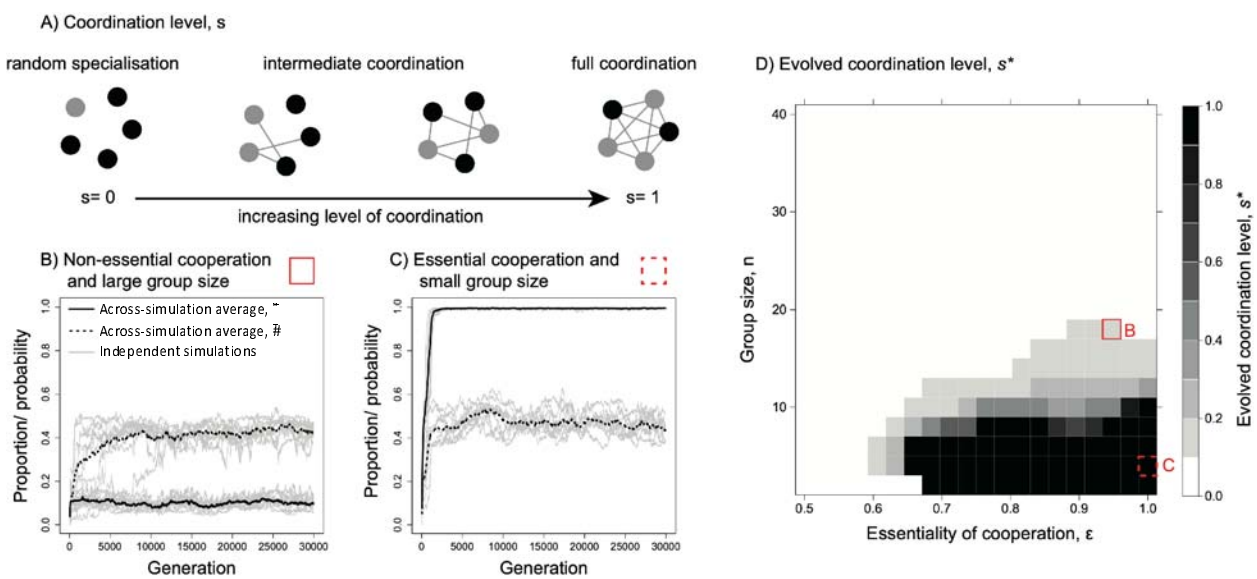


**Figure 4. The optimal level of coordination.** (A) If the level of coordination may evolve, then a spectrum of possible mechanisms could arise. At one extreme ($s = 0$), individuals may not coordinate at all (random specialisation). At the other extreme ($s = 1$), all individuals coordinate with one another and the variance in the realised proportion of helpers is minimised (fully

coordinated specialisation). In between ($0 < s < 1$), intermediate forms of coordination may arise in which not all individuals coordinate with one another and sub-optimal proportions of helpers can occur. (B & C) We present the across-simulation average level of coordination ($\bar{s}$, solid line) and target proportion of helpers ($\bar{t}$, dashed line) that evolves. We also plot individual simulation time series for the trait values (grey lines), where each line corresponds to a population average. We find that the traits converge stably within hundreds of generations. Moreover, the evolutionary outcomes (estimated by $s^*$ and $t^*$, see supplementary section G.1 for more details), depend on demographic and environmental conditions: panel B corresponds to the middle red box in panel D, whereas panel C to the red box at the lower right corner. Separate simulations have shown that these results are independent of starting conditions. (D) Smaller groups (lower $n$) and more essential cooperation (higher $\epsilon$) favour the evolution of fully coordinated specialisation ($s^* \approx 1$). Intermediate coordination ($0 < s^* < 1$) may evolve in more moderate conditions. (See figure S5 for the optimal target proportion of helpers, $t^*$). We modelled the cost of coordination as the increasing and decelerating function $c_n = \alpha\left(1 - e^{-\beta sn(n-1)/2}\right)$, where $sn\,(n-1)/2$ is the expected number of cell-to-cell interactions in a coordinated group with level of coordination $s$, and $0 < \alpha < 1$ and $\beta > 0$ are parameters that affect the scale and the shape of the cost of coordination curve, respectively. Here, $\beta = 0.01$ and $\alpha = 0.1$.

Overall, these predictions highlight that the level of coordination within a labour-dividing species is a trait that can be shaped by natural selection, and that fully coordinated specialisation is a mechanism that can be expected to evolve via intermediate forms in many cases. We also found that the degree to which intermediate levels of coordination may arise depends on the shape of the costs of coordination (see supplementary section G.2 and Figure S6). In particular, if there is a large up-front cost to coordination, then intermediate coordination ($0 < s^* < 1$) is less likely to arise—random specialisation will either be favoured (no coordination; $s^* \approx 0$) or systems will tend to evolve to fully coordinated specialisation ($s^* \approx 1$). Among microorganisms that employ coordination to divide labour, both cyanobacterial filaments and *Volvox carteri* colonies have very precise allocations of labour (full coordination) [50,51]. In contrast, the exact level of coordination in fruiting bodies such as *D. discoideum* and *Myxococcus xanthus* is less well known [52–54]. This highlights the need for further empirical work to determine the extent to which groups are “fully” coordinated in specific species.

Distribution of labour dividing mechanisms in the natural world

Most previous work on reproductive division of labour has tended to be either mechanistic, focusing on how different phenotypes are produced (caste determination), or evolutionary, focusing on why division of labour is favoured in the first place [1,6,23,36,39,40,55–59,14–21]. We have examined how evolutionary theory can also be used to explain why different mechanisms are used in different species [4,6,14–17]. In particular, we have shown that coordinated specialisation is more likely to be favoured over random specialisation in small groups, when coordination costs are low, and when there are larger fitness costs to deviating from optimal caste ratios. Our results identify social and environmental factors that could explain the distribution of mechanisms for division of labour that have been observed in bacteria and other microbes. Consistent with our predictions, coordination appears to be more prevalent in species where group sizes are smaller, cooperation is more essential, and/or the cost of coordination is lower (Table 1; see also supplementary section H). However, these connections are speculative, as more data is required for a formal comparative study across species. Our results also suggest a hypothesis for why random caste determination has not been widely observed in animal societies: during the initial evolution of complex animal societies, group sizes were likely to be small and costs of coordination might have been minor compared to each individual's day-to-day organismal metabolic expenditure.

| Species | Group size | Essentiality of cooperation | Cost of coordination | Type of specialisation |
|---|---|---|---|---|
| *Salmonella enterica* | Large | Low | High | Random |
| *Bacilus subtilis* | Medium | Medium | High | Random |
| *Volvox carteri* | Large | High | Low | Coordinated* |
| *Myxococcus xanthus* | Large | Medium/High | Medium/High | Coordinated |
| *Dictyostelium discoideum* | Large | Medium/High | Medium | Coordinated |
| *Anabaena cylindrica* | Small | Medium/High | Low | Coordinated |

**Table 1: Division of labour in microbes.** Examples of division of labour in bacteria and other microbes where the mechanism of specialisation has been identified [4,5,11,13,35,37,39,45,53,60,61]. For illustrative purposes, we provide approximate categorical estimates of relative group size, essentiality of cooperation, and the cost of coordination. This suggests roughly that coordinated specialisation has evolved most frequently in species with smaller groups, low coordination costs, and more essential cooperation. An across-species phylogeny-based comparative study, with a larger species sampling, would be required to formally test the predictions of our models [62]. This emphasizes the need for more data on the mechanisms underlying division of labour in these and other social microbe species (see supplementary section H). For example, persistence specialisation occurs randomly in large groups, and it has been suggested that this is cooperative [63]. **V. carteri* has a very precise allocation of labour[13], however there is observed variation in the number of germ and somatic cells in many Volvox species and so this may be a case where both coordinated specialisation and some random specialisation occur[11].

## Acknowledgements

We thank Anna Dewar, Kevin Foster, Andy Gardner, Ashleigh Griffin, Asher Leeks, Matishalin Patel, Tom Scott, and Daniel Unterwegger for their helpful comments and suggestions, and St. John's College, Oxford (GAC), the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program (grant ANR-17-EURE-0010; JP) and the ERC (Horizon 2020 Advanced Grant; SAW) for funding. The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC; http://dx.doi.org/10.5281/zenodo.22558) facility in carrying out this work.

## Data accessibility

All simulated data was generated using C and Matlab. The codes and generated data used for this study are available at:

https://github.com/mingpapilio/Codes_DOL_Mechanisms.

## Author Contributions

G.A.C., S.A.W., and J.P. conceived the study. G.A.C. and J.P. and designed and analysed the analytical models, G.A.C. and M.L. designed and analysed the simulation models. G.A.C. and S.A.W. wrote the first draft. All authors contributed toward writing the final manuscript.

## Declaration of Interests

The authors declare no competing interests.

## Supplementary information:

**A. Overview**

**B. Labour dividers and their fitness**

**C. Linear public goods**

**D. Alternative modelling assumptions**

**E. Alternative forms of cooperation**

**F. Risk and insurance for random specialisers**

**G. The optimal level of coordination**

**H. Estimating the characteristics of social microbes**

**Figure S1. Alternative modelling assumptions**

**Figure S2. Non-linear public goods**

**Figure S3. Alternative biological scenarios**

**Figure S4. Risk and insurance for random specialisers**

**Figure S5. Optimal target proportion of helpers**

**Figure S6. The shape of the cost of coordination**

# References:


1. Linksvayer, T. A. *et al.* Larval and nurse worker control of developmental plasticity and the evolution of honey bee queen-worker dimorphism. *J. Evol. Biol.* **24**, 1939–1948 (2011).
2. Flores, E. & Herrero, A. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat. Rev. Microbiol.* **8**, 39–50 (2010).
3. Weaver, N. Rearing of Honeybee Larvae on Royal Jelly in the Laboratory. *Science (80-. ).* **121**, 509–510 (1955).
4. Ackermann, M. *et al.* Self-destructive cooperation mediated by phenotypic noise. *Nature* **454**, 987–990 (2008).
5. Diard, M. *et al.* Stabilization of cooperative virulence by the expression of an avirulent phenotype. *Nature* **494**, 353–356 (2013).
6. Schwander, T., Lo, N., Beekman, M., Oldroyd, B. P. & Keller, L. Nature versus nurture in social insect caste differentiation. *Trends Ecol. Evol.* **25**, 275–282 (2010).
7. Hughes, W. O. H., Sumner, S., Van Borm, S. & Boomsma, J. J. Worker caste polymorphism has a genetic basis in Acromyrmex leaf-cutting ants. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9394–9397 (2003).
8. Anderson, K. E., Linksvayer, T. a & Smith, C. R. The causes and consequences of genetic caste determination in ants (Hymenoptera□: Formicidae). *Myrmecol. News* **11**, 119–132 (2008).
9. Jaffe, R., Kronauer, D. J. C., Kraus, F. B., Boomsma, J. J. & Moritz, R. F. A. Worker caste determination in the army ant Eciton burchellii. *Biol. Lett.* **3**, 513–516 (2007).
10. Zhang, Z. *et al.* Antibiotic production in Streptomyces is organized by a division of labor through terminal genomic differentiation. *Sci. Adv.* **6**, eaay5781 (2020).
11. Shelton, D. E., Desnitskiy, A. & Michod, R. E. Distributions of reproductive and somatic cell numbers in diverse Volvox (Chlorophyta) species. *Evol. Ecol. Res.* **14**, 707–727 (2012).
12. Griesemer, J. The Units of Evolutionary Transition. **1**, 67–80 (2000).
13. Smith, G. M. *A comparative study of the species of Volvox*. *Transactions of the American Microscopical Society* **63**, (1944).
14. Wahl, L. M. Evolving the division of labour: Generalists, specialists and task allocation. *J. Theor. Biol.* **219**, 371–388 (2002).
15. Wahl, L. M. The Division of Labor: Genotypic versus Phenotypic Specialization. *Am. Nat.* **160**, 135–145 (2002).
16. West, S. A. & Cooper, G. A. Division of labour in microorganisms: An evolutionary perspective. *Nat. Rev. Microbiol.* **14**, 716–723 (2016).
17. Duarte, A., Weissing, F. J., Pen, I. & Keller, L. An Evolutionary Perspective on Self-Organized Division of Labor in Social Insects. *Annu. Rev. Ecol. Evol. Syst.* **42**, 91–110 (2011).
18. Oster, G. F. & Wilson, E. O. *Caste and ecology in the social insects*. *Caste and ecology in the social insects* (Princeton University Press, 1978).
19. Wilson, E. O. The ergonomics of caste in the social insects. *Am. Econ. Rev.* **102**, 41–66 (1978).
20. Duarte, A., Pen, I., Keller, L. & Weissing, F. J. Evolution of self-organized division

of labor in a response threshold model. *Behav. Ecol. Sociobiol.* **66**, 947–957 (2012).

21. Duarte, A., Scholtens, E. & Weissing, F. J. Implications of behavioral architecture for the evolution of self-organized division of labor. *PLoS Comput. Biol.* **8**, e1002430 (2012).
22. Waibel, M., Floreano, D., Magnenat, S. & Keller, L. Division of labour and colony efficiency in social insects: Effects of interactions between genetic architecture, colony kin structure and rate of perturbations. *Proc. R. Soc. B Biol. Sci.* **273**, 1815–1823 (2006).
23. Bourke, A. F. G. *Principles of Social Evolution*. *Principles of Social Evolution* (Oxford University Press, 2011). doi:10.1093/acprof:oso/9780199231157.001.0001
24. Smits, W. K., Kuipers, O. P. & Veening, J. W. Phenotypic variation in bacteria: The role of feedback regulation. *Nat. Rev. Microbiol.* **4**, 259–271 (2006).
25. Süel, G. M., Garcia-Ojalvo, J., Liberman, L. M. & Elowitz, M. B. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* **440**, 545–550 (2006).
26. Dubnau, D. & Losick, R. Bistability in bacteria. *Mol. Microbiol.* **61**, 564–572 (2006).
27. Balázsi, G., Van Oudenaarden, A. & Collins, J. J. Cellular decision making and biological noise: From microbes to mammals. *Cell* **144**, 910–925 (2011).
28. Delong, E. F. *et al.* Noise in Gene Expression Determines Cell Fate in Bacillus subtilis. *Science (80-. ).* **27**, 526–530 (2007).
29. Ackermann, M. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.* **13**, 497–508 (2015).
30. Parker, G. A. & Maynard Smith, J. Optimality theory in evolutionary biology. *Nature* **348**, 27–33 (1990).
31. May, R. M. Uses and Abuses of Mathematics in Biology. *Science (80-. ).* **303**, 790–793 (2004).
32. West, S. A., Diggle, S. P., Buckling, A., Gardner, A. & Griffin, A. S. The Social Lives of Microbes. *Annu. Rev. Ecol. Evol. Syst.* **38**, 53–77 (2007).
33. Frank, S. A. A general model of the public goods dilemma. *J. Evol. Biol.* **23**, 1245–1250 (2010).
34. West, S. A. & Buckling, A. Cooperation, virulence and siderophore production in bacterial parasites. *Proc. R. Soc. B Biol. Sci.* **270**, 37–44 (2003).
35. Veening, J.-W. *et al.* Transient heterogeneity in extracellular protease production by Bacillus subtilis. *Mol. Syst. Biol.* **4**, 184 (2008).
36. Cooper, G. A. & West, S. A. Division of labour and the evolution of extreme specialization. *Nat. Ecol. Evol.* **2**, 1161–1167 (2018).
37. Koufopanou, V. The Evolution of Soma in the Volvocales. *Am. Nat.* **143**, 907–931 (2002).
38. Maynard Smith, J. & Harper, D. *Animal signals*. (Oxford University Press, 2003).
39. Strmecki, L., Greene, D. M. & Pears, C. J. Developmental decisions in Dictyostelium discoideum. *Dev. Biol.* **284**, 25–36 (2005).
40. Rossetti, V., Schirrmeister, B. E., Bernasconi, M. V. & Bagheri, H. C. The evolutionary path to terminal differentiation and division of labor in cyanobacteria.

*J. Theor. Biol.* **262**, 23–24 (2010).
41. Inglis, R. F., Ryu, E., Asikhia, O., Strassmann, J. E. & Queller, D. C. Does high relatedness promote cheater-free multicellularity in synthetic lifecycles? *J. Evol. Biol.* **30**, 985–993 (2017).
42. West, S. A. & Herre, E. A. Stabilizing Selection and Variance in Fig Wasp Sex Ratios. *Evolution (N. Y).* **52**, 475–485 (1998).
43. Nagelkerke, C. J. & Sabelis, M. W. Precise control of sex allocation in pseudo-arrhenotokous phytoseiid mites. *J. Evol. Biol.* **11**, 649–684 (1998).
44. Green, R. F., Gordh, G. & Hawkins, B. A. Precise Sex Ratios in Highly Inbred Parasitic. *Am. Nat.* **120**, 653–665 (1982).
45. Flores, E. & Herrero, A. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat. Rev. Microbiol.* **8**, 39–50 (2010).
46. Dionisio, F. & Gordo, I. The Tragedy of the Commons, the Public Good Dilemma and the Meaning of Rivalry and Excludability in Evolutionary Biology. *Evol. Ecol. Res.* **8**, 321–332 (2006).
47. Gardner, A. & Grafen, A. Capturing the superorganism: A formal theory of group adaptation. *J. Evol. Biol.* **22**, 659–671 (2009).
48. Hamilton, W. D. The genetical evolution of social behaviour I-II. *J. Theor. Biol.* **7**, 1–52 (1964).
49. Peña, J., Nöldeke, G. & Lehmann, L. Evolutionary dynamics of collective action in spatially structured populations. *J. Theor. Biol.* **382**, 122–136 (2015).
50. Flores, E. & Herrero, A. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nature Reviews Microbiology* (2010). doi:10.1038/nrmicro2242
51. Solari, C. A., Kessler, J. O. & Goldstein, R. E. A General Allometric and Life-History Model for Cellular Differentiation in the Transition to Multicellularity. *Am. Nat.* **181**, 369–380 (2013).
52. Maree, A. F. M. & Hogeweg, P. How amoeboids self-organize into a fruiting body: Multicellular coordination in Dictyostelium discoideum. *Proc. Natl. Acad. Sci.* **98**, 3879–3883 (2002).
53. Velicer, G. J., Kroos, L. & Lenski, R. E. Developmental cheating in the social bacterium Myxococcus xanthus. *Nature* **404**, 598–601 (2000).
54. Madgwick, P. G., Stewart, B., Belcher, L. J., Thompson, C. R. L. & Wolf, J. B. Strategic investment explains patterns of cooperation and cheating in a microbe. *Proc. Natl. Acad. Sci.* **115**, E4823–E4832 (2018).
55. Kamakura, M. Royalactin induces queen differentiation in honeybees. *Nature* **473**, 478 (2011).
56. Gavrilets, S. Rapid transition towards the division of labor via evolution of developmental plasticity. *PLoS Comput. Biol.* **6**, e1000805 (2010).
57. Ispolatov, I., Ackermann, M. & Doebeli, M. Division of labour and the evolution of multicellularity. *Proc. R. Soc. B Biol. Sci.* **279**, 1768–1776 (2012).
58. Linksvayer, T. A. Ant species differences determined by epistasis between brood and worker genomes. *PLoS One* **2**, e994 (2007).
59. Walsh, J. T., Warner, M. R., Kase, A., Cushing, B. J. & Linksvayer, T. A. Ant nurse workers exhibit behavioural and transcriptomic signatures of specialization on larval stage. *Anim. Behav.* **141**, 161–169 (2018).

60. Herron, M. D., Hackett, J. D., Aylward, F. O. & Michod, R. E. Triassic origin and early radiation of multicellular volvocine algae. *Proc. Natl. Acad. Sci.* **106**, 3254–3258 (2009).
61. Strassmann, J. E., Zhu, Y. & Queller, D. C. Altruism and social cheating in the social amoeba Dictyostelium discoideum. *Nature* **408**, 965–967 (2000).
62. Fisher, R. M., Cornwallis, C. K. & West, S. A. Group formation, relatedness, and the evolution of multicellularity. *Curr. Biol.* **23**, 1120–1125 (2013).
63. Gardner, A., West, S. A. & Griffin, A. S. Is Bacterial Persistence a Social Trait? *PLoS One* **2**, e752 (2007).

# Mechanisms to divide labour in social microoganisms
# Supporting Information

Guy A. Cooper*[1,2], Jorge Peña†[3], Ming Liu‡[2], and Stuart A. West§[2]

[1]St. John's College, Oxford, United Kingdom
[2]Department of Zoology, University of Oxford, Oxford, United Kingdom
[3]Institute for Advanced Study in Toulouse, University of Toulouse Capitole, Toulouse, France

## Contents

**A Overview** **2**

**B Labour dividers and their fitness** **2**
B.1 Fully coordinated specialisation . . . 2
B.2 Random specialisation . . . 3

**C Linear public goods** **3**

**D Alternative modelling assumptions** **4**
D.1 Discrete proportions of helpers . . . 4
D.2 Optimal random specialisers . . . 4
D.3 Non-linear benefits to cooperation . . . 6

**E Alternative forms of cooperation** **6**
E.1 Reproductives benefit from the proportion of helpers . . . 6
E.2 Reproductives benefit from the number of helpers . . . 8
E.3 Reproductives benefit from the ratio of helpers to reproductives . . . 9

**F Risk and insurance for random specialisers** **11**

**G The optimal level of coordination** **13**
G.1 Individual based model . . . 13
G.1.1 Simulation initialisation . . . 13
G.1.2 Coordination network . . . 13
G.1.3 Individual specialisation . . . 13
G.1.4 Fitness calculation . . . 14
G.1.5 Creating a new generation . . . 14
G.1.6 Simulation results . . . 14
G.1.7 A note on our assumptions . . . 15
G.2 Alternate costs of coordination . . . 16
G.3 Heuristic analytical model . . . 16

**H Estimating the characteristics of social microbes** **18**

*E-mail address: guy.cooper@zoo.ox.ac.uk
†E-mail address: jorge.pena@iast.fr
‡E-mail address: ming.liu@zoo.ox.ac.uk
§E-mail address: stuart.west@zoo.ox.ac.uk

# A Overview

This manuscript is organized as follows. Section B defines the strategies of fully coordinated specialisation and random specialisation, and characterises their fitness. Section C derives the main analytical result of the main text (Equation 5), that is, the condition for fully coordinated specialisation to be favoured over random specialisation when the underlying division of labour is modelled as a linear public goods game. Section D re-examines key assumptions and approximations that simplified the analysis of the public goods model. Section E considers several models for alternative biological systems leading to different kinds of public goods. Both sections D and E demonstrate the robustness and generality of our results when relaxing our assumptions and approximations. Section F presents a numerical investigation on the impact of asymmetry in the fitness functions on the optimal probability of becoming a helper, and how this can mitigate some of the costs of random specialisation. Section G examines the optimal level of coordination using individual based simulations as well as an analytical model (as presented in Fig. 4 of the main text). Finally, section H presents details on how we categorised the relative group size, essentiality of cooperation, and cost of coordination for the different labour-dividing microbial species presented in Table 1 of the main text.

# B Labour dividers and their fitness

We assume that a single individual arrives on an empty patch and, through a fixed series of replications, forms a clonal group of $n$ individuals that is composed of $k$ sterile helpers and $n-k$ pure reproductives, where $k \in \mathcal{N} = \{0, 1, 2, \ldots, n\}$. The reproductive success (fitness) of the group, $g_{k,n}$, as measured by the per capita number of dispersing offspring at the end of the group life cycle, is given by

$$g_{k,n} = \frac{1}{n}(n-k) f_{k,n}, \tag{S1}$$

where $f_{k,n}$ is the fecundity of each reproductive in the group. We assume that $f_{k,n}$ depends only on the proportion of helpers in the group, $p = k/n$ with $p \in \mathcal{P} = \{0, 1/n, 2/n, \ldots, (n-1)/n, 1\}$, so that we can write $f_{k,n} = F(p)$, where $F$ is a real function that is increasing on the interval $[0, 1]$. This assumption allows us to rewrite (S1) as

$$g_{k,n} = (1 - k/n)\, F\,(k/n) = G(k/n), \tag{S2}$$

where we have defined

$$G(p) = (1-p)F(p). \tag{S3}$$

## B.1 Fully coordinated specialisation

With fully coordinated specialisation (C), we assume that some mechanism, such as signalling between cells, ensures that groups always form with the optimal proportion of helpers, $p^*$, where

$$p^* = \frac{k^*}{n}, \text{ with } k^* = \arg\max_{k \in \mathcal{N}} g_{k,n}. \tag{S4}$$

We posit that the fitness of a group of coordinated specialisers is given by:

$$w_{\mathrm{C}}(p^*) = (1 - c_n) \max_{k \in \mathcal{N}} g_{k,n} = (1 - c_n)\, G(p^*), \tag{S5}$$

where $c_n \in [0, 1]$ is the cost of coordination, which we assume is a non-decreasing function of the group size, $n$. A parsimonious choice, that we use to illustrate all of our results, is to set

$$c_n = \alpha(1 - e^{-\beta n(n-1)/2}), \tag{S6}$$

where $n(n-1)/2$ is the number of cell-to-cell interactions in a fully coordinated group, $\alpha \in [0, 1]$ controls the scale of the cost, and $\beta > 0$ determines how diminishing the cost is (as presented in Fig. 3 of the main text). Note, however, that our main result (Equation 5 in the main text, to be derived below) is independent of the particular functional form of the cost of coordination.

## B.2 Random specialisation

With random specialisation (R), each individual in the group independently becomes a helper with a given probability $q$, and a reproductive otherwise. Hence, the number of helpers in the group, that we denote by $K$, is a binomial random variable with parameters $n$ and $q$ (i.e., $K \sim \text{Binomial}(n, q)$). In the following it will also be convenient to write $Q = K/n$ for the random variable giving the proportion of helpers in the group. Within this framework, the expected fitness of a group of random specialisers is simply given by

$$w_{\mathrm{R}}(q) = \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} g_{k,n}. \tag{S7}$$

# C Linear public goods

Our main model assumes that the fecundity function $F$ is linear in $p$ and given by

$$F(p) = 1 - \epsilon + \epsilon p, \tag{S8}$$

where $\epsilon \in [0, 1]$ is a parameter measuring the "essentiality of cooperation". For $\epsilon = 0$, cooperation by helpers is totally unessential, as the fecundity of reproductives is constant and equal to 1. For $\epsilon = 1$, cooperation by helpers is totally essential and (S8) reduces to $F = p$, so that the fecundity of reproductives equals the proportion of helpers in the group. For $0 < \epsilon < 1$, the fecundity of reproductives is the sum of a constant term (i.e., $1 - \epsilon$) and a term proportional to the proportion of helpers in the group (i.e., $\epsilon p$).

Replacing (S8) into (S3) and simplifying we obtain

$$G(p) = (1-\epsilon)(1-p) + \epsilon p(1-p). \tag{S9}$$

To find the fitness of coordinated specialisers, we first assume that $p$ is a continuous variable, and calculate the derivative

$$G'(p) = 2\epsilon - 1 - 2\epsilon p.$$

This derivative is decreasing in $p$ (i.e., $G(p)$ is concave), and has a single root, $\hat{p}$, given by

$$\hat{p} = 1 - \frac{1}{2\epsilon}. \tag{S10}$$

Such a root lies in the interval $(0, 1)$ if and only if $1/2 < \epsilon \leq 1$. Otherwise the maximiser of $G(p)$ (and hence the optimal allocation of helpers) is given by $\hat{p} = 0$ (i.e., it is optimal to have no helpers). To avoid this trivial scenario without division of labour, henceforth we assume that $1/2 < \epsilon \leq 1$ holds. Further, to make progress we approximate the optimal allocation of helpers, $p^*$, by $\hat{p}$. The actual optimal value $p^*$ will be a value near $\hat{p}$ but constrained by the permissible group compositions, since $p^* \in \mathcal{P}$ (cf. section D.1). Note that the approximate optimal proportion (S10) is independent of $n$, and that it is an increasing function of $\epsilon$ such that $\hat{p} = 1/2$ when $\epsilon = 1$. An approximation to the fitness of fully coordinated specialisers can then be obtained by letting $p^* \approx \hat{p}$ in equation (S5), that is, by letting

$$w_{\mathrm{C}}(p^*) \approx (1 - c_n)\, G(\hat{p}). \tag{S11}$$

To find the fitness of random specialisers, we replace (S9) into (S7) and simplify to obtain

$$\begin{aligned} w_{\mathrm{R}}(q) &= \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} (1-k/n)(1-\epsilon+\epsilon k/n) \\ &= (1-\epsilon)\mathrm{E}\left[(1-K/n)\right] + \epsilon \mathrm{E}\left[(1-K/n)K/n\right] \\ &= (1-\epsilon)(1-q) + \epsilon q(1-q) - \epsilon \frac{q(1-q)}{n} \qquad \text{(S12)} \\ &= G(q) - \epsilon \mathrm{Var}(Q), \qquad \text{(S13)} \end{aligned}$$

where we have made use of the first two moments of the binomial distribution, $\mathrm{E}[K] = nq$, $\mathrm{E}[K^2] = nq(1-q) + (nq)^2$, of expression (S9), and of the fact that $\mathrm{Var}(Q) = q(1-q)/n$.

In order to determine the condition under which coordinated specialisation is favoured over random specialisation, we assume in a first step that random specialisers play the strategy $q = p^*$, so that their fitness is given by

$$w_{\mathrm{R}}(p^*) = G(p^*) - \epsilon \mathrm{Var}(P^*), \tag{S14}$$

where $P^* = K^*/n$ and $K^* \sim \mathrm{Binomial}(n, p^*)$ (and hence $\mathrm{Var}(P^*) = p^*(1-p^*)/n$). This assumption simplifies our calculations and leads to results that are qualitatively similar to those that arise from the more parsimonious assumption that random specialisers play the strategy that maximises their fitness (cf. section D.2). Comparing expressions (S5) and (S14), it follows that the condition $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(p^*)$ can be written as

$$c_n < c^*, \tag{S15}$$

where the critical coordination cost, $c^*$, is given by

$$c^* = \frac{\epsilon}{G(p^*)} \mathrm{Var}\left(P^*\right). \tag{S16}$$

Equation (S16) makes it explicit the fact that the critical cost of coordination may be decomposed into a measure of the deviation from the optimal allocation of labour, $\mathrm{Var}\left(P^*\right)$, and a quantity that captures the relative cost of deviating from the optimal proportion of helpers, $\epsilon/G(p^*)$. Furthermore, since $G(p^*)$ is (approximately) independent of $n$, we can say that the effect of increasing the group size acts primarily on the deviation of groups from the optimal proportion of helpers, $\mathrm{Var}\left(P^*\right)$. In contrast, a larger essentiality of cooperation both (i) pushes $p^*$ closer to $1/2$ (which in turns increases the variance $\mathrm{Var}(P^*) = (2\epsilon - 1)/(4n\epsilon^2)$) and (ii) increases the cost of deviation (larger $\epsilon/G(p^*) = 4\epsilon^2$) and thus acts via both factors.

To obtain a simple expression of $c^*$ in terms of our parameters $n$ and $\epsilon$, we approximate $p^*$ by $\hat{p}$ as given in (S10) to obtain

$$c^* \approx \frac{2\epsilon - 1}{n}, \tag{S17}$$

which is increasing in the essentiality of cooperation $\epsilon$ and decreasing in group size $n$. Expression (S15) together with the approximation yields expression 5 in the main text.

# D Alternative modelling assumptions

## D.1 Discrete proportions of helpers

In section C, we approximated the optimal proportion of helpers in coordinated groups as a continuous variable $p^* \approx \hat{p} \in [0, 1]$. In reality this quantity is discrete as it is given by $p^* = k^*/n$ and there can only be an integer number of helpers in the group ($k^* \in \mathcal{N}$, so that $p^* \in \mathcal{P}$).

Here, we show that this assumption is relatively innocuous. First, it is clear that $p^*$ tends to $\hat{p}$ as $n$ grows large (Fig. S1A). Second, even for relatively small group size $n$ the predictions of the model are relatively independent of this assumption. To show this, we repeat our analysis without making the continuous approximation. That is, we evaluate the conditions $c_n < c^*$ (coordinated specialisation favoured) and $c_n > c^*$ (random specialisation favoured), where $c^*$ is given by (S16) and where $p^*$ is not approximated as $\hat{p}$ but is calculated numerically for each combination of parameters considered, so that $p^* \in \mathcal{P}$. The results of this analysis are shown in Fig. S1B. As expected, we find that the parameter space boundary is more jagged but that the broad results are similar to those of the continuous treatment.

## D.2 Optimal random specialisers

In section C, we assumed that the probability of becoming a helper, $q$, was approximately equal to the optimal proportion of helpers (i.e., we let $q = \hat{p} \approx p^*$) (Fig. S1A). A more realistic assumption is that the probability of being a helper in random specialisers is the one optimising fitness, so that random specialisers play strategy $q = q^*$, with

$$q^* = \arg\max_{0 \leq q \leq 1} w_{\mathrm{R}}(q). \tag{S18}$$

This alternative modelling assumption makes the fitness of random specialisers be larger (as $w_{\mathrm{R}}(q^*) \geq w_{\mathrm{R}}(p^*)$ necessarily holds) and thus will make, all else being equal, random specialisation more likely to be favoured over coordinated specialisation.

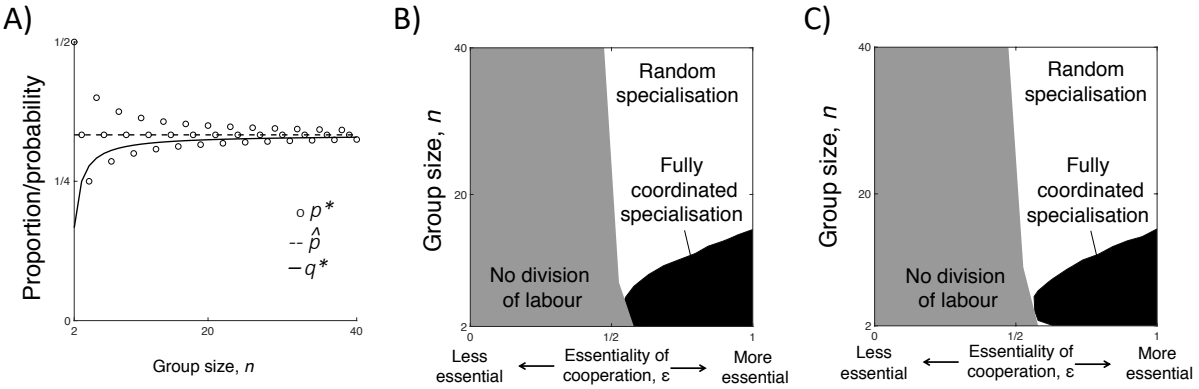


Figure S1: **Alternative modelling assumptions**. A) In the model presented in the main text, we assumed that the probability that random specialisers adopt a helper role was equal to the optimal proportion of helpers in the group ($p^*$) as opposed to the probability that optimises fitness for random specialisers ($q^*$). In turn, we approximated this optimal proportion of helpers by a continuous variable $\hat{p}$. Here we show that all three of these quantities, $p^*$ (open circles), $\hat{p}$ (dashed line), and $q^*$ (solid line) converge to the same value as group size ($n$) increases. Here, $\epsilon = 3/4$, leading to $\hat{p} = (1 - 2\epsilon)/(2\epsilon)$. B) We evaluate the conditions $c_n < c^*$ (coordinated specialisation favoured) or $c_n > c^*$ (random specialisation favoured), without approximating $p^*$ by $\hat{p}$ (as in the main text). Instead, we evaluate the optimal value $p^*$ numerically for each parameter combination considered. C) We evaluate the condition $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(q^*)$ (coordinated specialisation favoured) or $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(q^*)$ (random specialisation favoured), while allowing for the probability that random specialisers adopt a helper role to evolve to the value that maximises the fitness of random specialisers ($q^*$). In both scenarios B) and C), we find that smaller group sizes (smaller $n$) and more essential cooperation (higher $\epsilon$) favour coordinated specialisation, in agreement with the results presented in the main text. The cost of coordination is given by (S6) with $\alpha = 0.1$ and $\beta = 0.01$. The red area occurs where division of labour is not favoured, that is when the optimal proportion of helpers is zero.

To find $q^*$ for the linear public goods model, we take the derivative of (S12) with respect to $q$:

$$w'_{\mathrm{R}}(q) = 2\epsilon - 1 - \epsilon/n - 2\epsilon(n-1)q/n. \tag{S19}$$

This derivative is decreasing in $q$ (hence $w_{\mathrm{R}}(q)$ is concave) and has a single root $q^*$ given by

$$q^* = 1 + \frac{1}{2(n-1)} - \frac{n}{2\epsilon(n-1)}. \tag{S20}$$

Such root lies in the interval $(0,1)$ if $w'_{\mathrm{R}}(0) = (2n\epsilon - n - \epsilon)/n > 0$ holds, or equivalently, if $\epsilon > n/(2n-1)$ holds, which we assume in the following. Expression (S20) is increasing in both $n$ and $\epsilon$. We also note that

$$\hat{p} - q^* = \frac{1-\epsilon}{2(n-1)\epsilon} > 0, \tag{S21}$$

and hence (i) $\hat{p}$ always overestimates $q^*$, but that (ii) the difference between the two values is inversely proportional to $n$ and goes to zero as $n$ grows large. Thus, $\hat{p}$ approximates $q^*$ relatively well for relatively large $n$, which justifies our use of $\hat{p}$ as an approximation in section C (see Fig. S1A).

We further note that the derivative (S19) can be understood as a selection gradient on $q$. As such selection gradient is decreasing in $q$, the point $q^*$ is not only a fitness maximum, but also the value to which evolution by small-step mutations would eventually approach. In other words, $q^*$ is a convergence stable strategy [1, 2].

We present numerical results in Fig. S1C, solving for the optimal mechanism by determining the conditions under which either $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(q^*)$ (coordinated specialisation favoured) or $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(q^*)$ (random specialisation favoured) holds. Overall, we find similar qualitative results hold as in the simplified model.

### D.3 Non-linear benefits to cooperation

In the main model, we assumed that reproductive fecundity depends linearly on the proportion of helpers in the group. Here, we consider the possibility that there exists a non-linear dependence on the proportion of helpers by modelling reproductive fecundity with the following sequence:

$$f_{k,n} = 1 - \epsilon + \epsilon\,(k/n)^{\lambda}, \tag{S22}$$

where $\lambda > 0$ is a parameter controlling the shape of the return from an increasing proportion of helpers. When $\lambda < 1$, there is a large initial return from the addition of the first helper, followed by a decelerating rate of return as the proportion of helpers increases (i.e., $f_{k,n}$ is concave; see Fig. S2A). When $\lambda > 1$ there is a negligible initial return from the addition of the first helper followed by an accelerating rate of return as the proportion of helpers increases (i.e., $f_{k,n}$ is convex; see Fig. S2B). In all cases (S22) is monotonically increasing from zero to one.

We considered the parameter discretisation, $n \in \{2, 4, \ldots, 38, 40\}$ and $\epsilon \in \{0.1, 0.1 + 1 \times 0.9/39, \ldots, 0.1 + 38 \times 0.9/39, 1\}$ and for each combination of parameters, we solved numerically for the optimal proportion of helpers, $p^*$, that maximises $G(p)$. For each combination of parameters, we determined the optimal mechanism (i.e., whether $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(p^*)$ or $w_{\mathrm{C}}(p^*) < w_{\mathrm{R}}(p^*)$ holds), to ascertain whether accelerating ($\lambda > 1$), or decelerating ($\lambda < 1$) returns affect the predictions of our analysis. We find that the same broad results hold as in the linear public goods game (see Fig. S2C and S2D). We further find that decelerating benefits can lead to fully coordinated specialisation being favoured even for relatively low essentiality of cooperation (Fig. S2C) but that accelerating benefits leads to fully coordinated specialisation being somewhat more advantageous than decelerating benefits at higher group sizes (Fig. S2D).

## E Alternative forms of cooperation

In the main model, we assumed that the fecundity of reproductives depends on the proportion of helpers in the group, $k/n$. Here, we consider the conceptual underpinnings of this assumption (section E.1) and analyse alternative scenarios where reproductive fecundity depends on the number of helpers in the group, $k$ (section E.2), or on the ratio of helpers to reproductives, $k/(n-k)$ (section E.3).

### E.1 Reproductives benefit from the proportion of helpers

There are two key scenarios where reproductive fecundity can be modelled as depending on the proportion of helpers in the group, $k/n$.

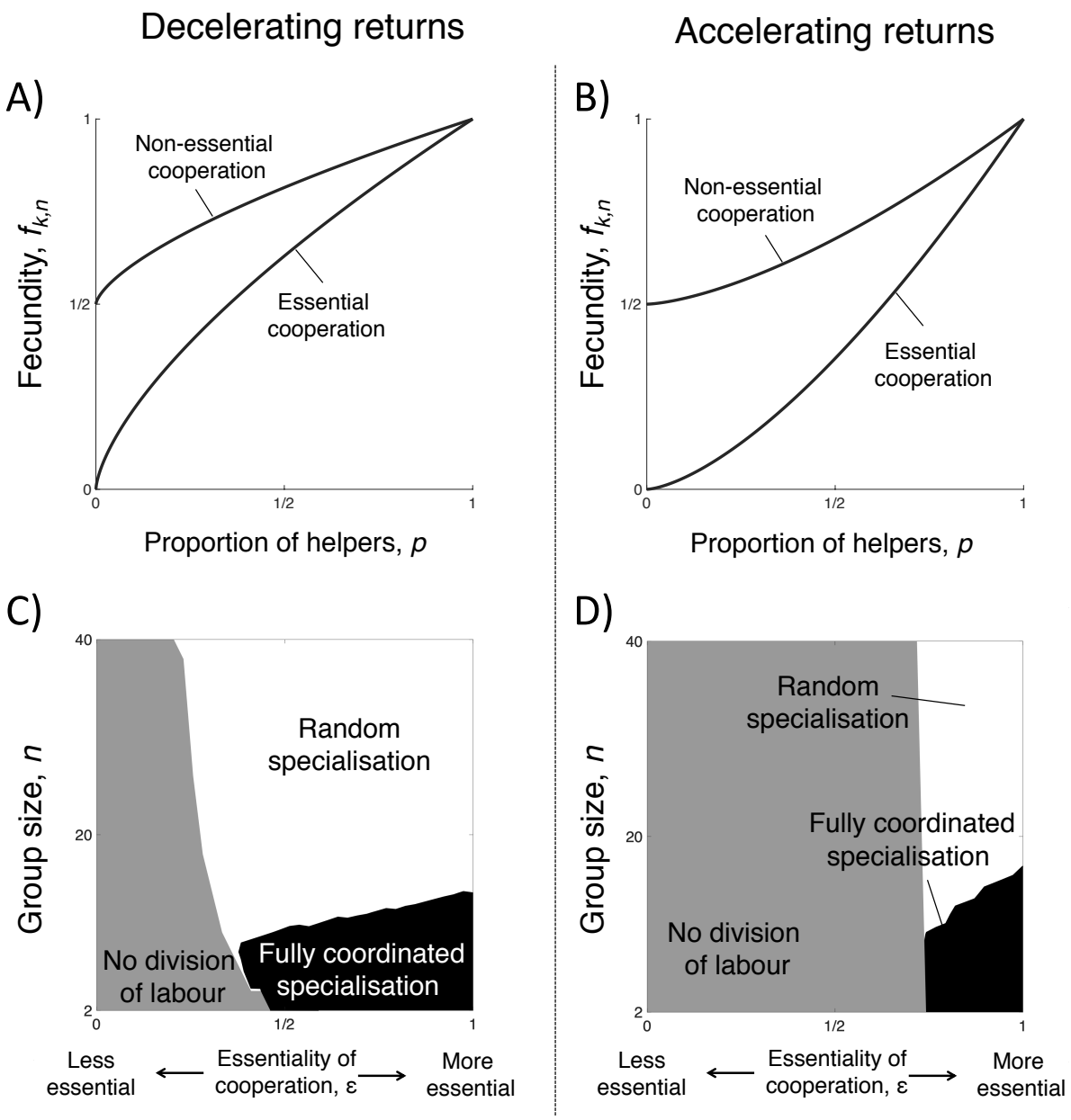


Figure S2: **Non-linear public goods**. We consider an extension of the model allowing the fecundity of reproductives to depend non-linearly on the proportion of helpers in the group. A) If $\lambda < 1$, then the presence of helpers provides an initially large increase in the fecundity of reproductives but then only provides benefits at a decreasing rate as the proportion of helpers goes up (decelerating cooperation; here $\lambda = 2/3$). B) If $\lambda > 1$, then the presence of helpers provides no initial increase in the fecundity of reproductives but then provides benefits at an increasing rate as the proportion of helpers goes up (accelerating cooperation; here $\lambda = 3/2$). In both A) and B) we have set $\epsilon = 1/2$ (non-essential cooperation) and $\epsilon = 1$ (essential cooperation). C) and D) We numerically evaluate the conditions in which $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(p^*)$ (coordinated specialisation favoured) or $w_{\mathrm{C}}(p^*) < w_{\mathrm{R}}(p^*)$ (random specialisation favoured) holds while assuming either that $\lambda = 2/3$ (decelerating cooperation; panel C). or $\lambda = 3/2$ (accelerating cooperation; panel D). The cost of coordination is given by (S6) with $\alpha = 0.1$ and $\beta = 0.01$. In both scenarios, we find that small group sizes (lower $n$) and more essential cooperation (higher $\epsilon$) favour coordinated specialisation. The red area occurs where division of labour is not favoured, that is when the optimal proportion of helpers is zero.

First, the benefits provided by helpers may constitute a good that is consumed by its beneficiaries (sometimes termed a congestible or rivalrous good), wherein the benefit experienced by one individual proportionally decreases the amount of good that may benefit its neighbours [3, 4]. For instance, the secretion of an extracellular product that must be absorbed and digested in order to provide benefits (as happens in populations of *B. subtilis* and *A. cylindrica*) can be conceptualised as such a good [5, 6, 7]. In particular, if the good is non-excludible (i.e., all individuals use the benefits) then the benefits conferred to each reproductive is the total amount of help (in the linear case, approximated as the number of helpers, $k$) divided by the number of individuals in the group ($n$; reproductives and helpers alike) [3]. That helpers also partake in the consumption of the public good may be considered a type of soaking [8].

Second, the amount of good that each helper provides may depend on the number of individuals in the group such that, all else being equal, larger groups entail less of a benefit to the group. For instance, helpers in *V. carteri* beat their flagella to keep the colony afloat at the optimal height in the water column for photosynthesis [9, 10]. In this case, the contribution of each helper (the degree to which it helps keep the colony at the right height) depends inversely on the size of the group as larger colonies are more difficult to keep afloat.

## E.2 Reproductives benefit from the number of helpers

An alternative assumption is that the fecundity of reproductives depends on the number of helpers in the group, $k$. This would imply that the contribution of each helper to the total good does not depend on the size of the group and that the benefit conferred to one individual does not decrease the benefits available for another. The benefits provided by helpers in *S. enterica* infections is an example of such a good (a non-congestible or non-rivalrous good) [3, 4]. In this case, helpers trigger a host immune response that wipes out competing microbial strains so that reproductives may then proliferate without competition [11, 12]. Thus, the competitive advantage afforded by the host immune response is not used up or depleted by any of its beneficiaries, unless there are so few niches vacated that reproductives then compete with one another.

We would like to know whether we recover similar results as the ones we obtained assuming that reproductives benefit from the proportion of helpers in an alternative biological model in which the fecundity of reproductives depends on the number of helpers in the group, $k$. To this end we posit the following specific sequence for reproductive fecundity:

$$f_{k,n} = 1 - \epsilon + \epsilon k.$$

Replacing this expression into (S3) and by a slight abuse of notation (as $G$ now depends explicitly on the group size $n$) we obtain

$$G(p) = (1-p)(1-\epsilon+\epsilon np). \tag{S23}$$

Treating $p$ as a continuous variable, we calculate the first derivative of $G$ with respect to $p$ as

$$G'(p) = n\epsilon - (1-\epsilon) - 2\epsilon np.$$

This derivative is decreasing in $p$ (i.e., $G(p)$ is concave), and has a single root $\hat{p}$ given by

$$\hat{p} = \frac{1}{2} - \frac{1-\epsilon}{2n\epsilon}. \tag{S24}$$

Such root lies in the interval $(0,1)$ if $\epsilon > 1/(n+1)$, which we assume henceforth.

Using arguments similar to the ones we used in section C, it follows that we can approximate the fitness of a coordinated group as

$$w_{\mathrm{C}}(p^*) \approx (1-c_n)G(\hat{p}),$$

while the fitness of a randomly specialising group randomising with probability $q$ can be written as

$$\begin{aligned} w_{\mathrm{R}}(q) &= \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} (1-k/n)(1-\epsilon+\epsilon k) \\ &= G(q) - \epsilon q(1-q) \\ &= G(q) - \epsilon n \mathrm{Var}(Q). \end{aligned}$$

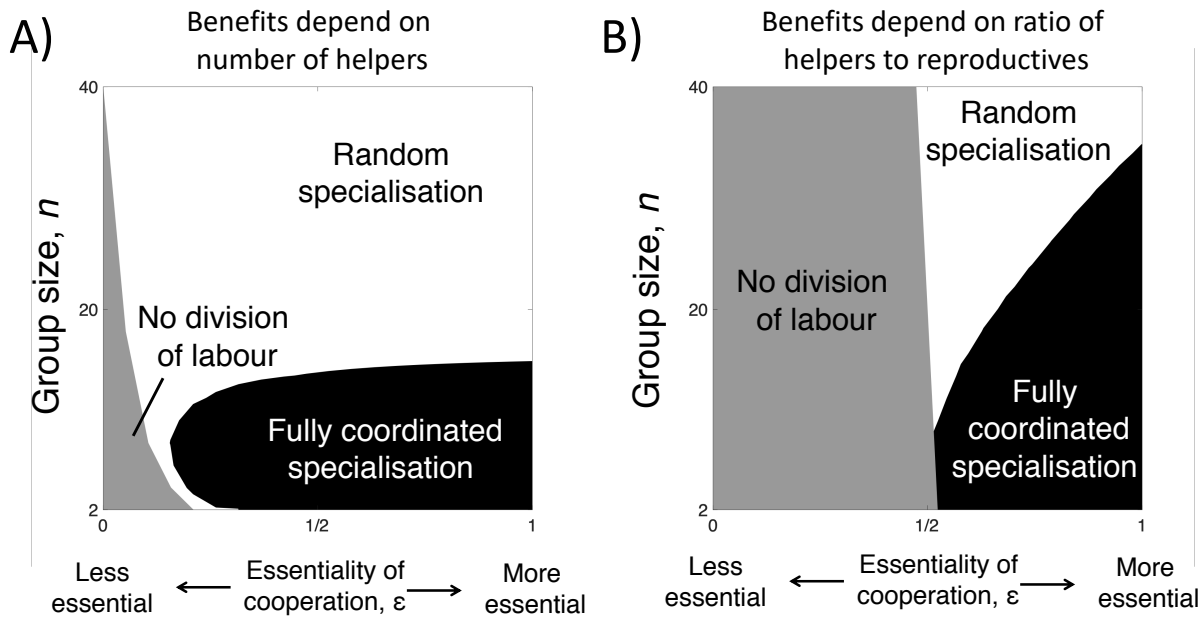


Figure S3: **Alternative biological scenarios**. A) We consider a scenario where the fecundity of reproductives depends on the number of helpers in the group, evaluating when $c_n < c^*$ (coordinated specialisation favoured) or $c_n > c^*$ (random specialisation favoured) holds, and where $c^*$ is calculated according to equation (S25). B) We consider a scenario where the fecundity of reproductives depends on the ratio of helpers to reproductives, evaluating when $c_n < c^*$ (coordinated specialisation favoured) or $c_n > c^*$ (random specialisation favoured) holds, and where $c^*$ is calculated with equation (S31). The cost of coordination is given by equation (S6) with $\alpha = 0.1$ and $\beta = 0.01$. In both cases, we find the same qualitative results as in the main analysis. That is, smaller group sizes (smaller $n$) and more essential cooperation (higher $\epsilon$) favour coordinated specialisation. The red area occurs where division of labour is not favoured, that is when the optimal proportion of helpers is zero.

Assuming that random specialisers play $q = p^*$, division of labour by coordinated specialisation is favoured over random specialisation if $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(p^*)$ holds. This condition can be written as $c_n < c^*$, where the critical cost $c^*$ is given by

$$c^* = \frac{\epsilon n \mathrm{Var}(P^*)}{G(p^*)} = \frac{\epsilon p^*}{1 - \epsilon + \epsilon n p^*} \approx \frac{n\epsilon - (1-\epsilon)}{n[(n-1)\epsilon + 1]}, \tag{S25}$$

where the approximation follows from assuming $p^* \approx \hat{p}$ with $\hat{p}$ given by (S24).

The critical cost $c^*$ is increasing in $\epsilon$, and decreasing in $n$ for relatively large $n$. This can be easily seen by taking the first forward difference of $c^*(n)$,

$$\Delta_n c^* = c^*(n+1) - c^*(n) = \frac{1 - \epsilon + 2n\epsilon - 3n\epsilon^2 - n^2\epsilon^2}{n(n+1)(n\epsilon+1)(1-\epsilon+n\epsilon)},$$

and realizing that this expression is negative for large $n$.

We present some numerical results in Fig. S3A, evaluating when $c_n < c^*$ (i.e., coordinated specialisation is favoured) or $c_n > c^*$ (i.e., random specialisation is favoured), where $c^*$ is given by (S25). As in the main model where reproductives benefit from the proportion of helpers, we find that smaller group sizes (small $n$) and more essential cooperation (higher $\epsilon$) favour coordinated specialisation.

## E.3 Reproductives benefit from the ratio of helpers to reproductives

In a second alternative scenario, only reproductive individuals benefit from the collective good. In this case, reproductive fecundity is determined by the number of helpers (approximate amount of collective good) divided

by the number of reproductives (the number of beneficiaries of the good). This gives the following expression for the fecundity of a reproductive in a group of size $n$ with $k$ helpers:

$$f_{k,n} = \left(1 - \epsilon + \epsilon \frac{k}{n-k}\right)[k < n], \tag{S26}$$

where $[k < n]$ is an Iverson bracket indicating that $f_{k,n} = 0$ if $k = n$.

Replacing this expression into (S1) and simplifying, we obtain the following formula for the fitness of a group of size $n$ with $k$ helpers:

$$g_{k,n} = \left(1 - \epsilon + \frac{(2\epsilon - 1)k}{n}\right)[k < n]. \tag{S27}$$

This sequence is unimodal (first increasing, then decreasing) in $k$ if $\epsilon > 1/2$ holds, which we assume henceforth. In this case, $k^* = n - 1$ maximises $g_{k,n}$ for all $n$, so that the optimal proportion of helpers is given by $p^* = (n-1)/n$. It follows that the fitness for coordinated specialisers is given by

$$w_{\mathrm{C}}(p^*) = (1 - c_n) g_{k^*,n} = (1 - c_n) g_{n-1,n} = (1 - c_n)\frac{(n-2)\epsilon + 1}{n}. \tag{S28}$$

Contrastingly, the expected fitness for random specialisers is given by:

$$\begin{aligned} w_{\mathrm{R}}(q) &= \sum_{k=0}^{n-1} \binom{n}{k} q^k (1-q)^{n-k} [1 - \epsilon + (2\epsilon - 1)k/n] \\ &= (1-\epsilon)(1-q^n) + (2\epsilon - 1)(q - q^n) \\ &= 1 - \epsilon + (2\epsilon - 1)q - \epsilon q^n. \end{aligned} \tag{S29}$$

At this point we deviate from our previous approximations and compare the fitness of coordinated specialisers and the fitness of random specialisers by evaluating the condition $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(q^*)$, where $q^*$ is the optimal probability of becoming a helper for random specialisers. We do this because this model leads to highly asymmetric group fitness $g_{k,n}$, which select for random specialisers to strongly under-weight the probability of becoming a helper ($q^* \ll p^*$; see section F) and therefore the approximation $q^* \approx p^*$ is less accurate.

The derivative of $w_{\mathrm{R}}(q)$ (S29) with respect to $q$ is given by $w'_{\mathrm{R}}(q) = 2\epsilon - 1 - n\epsilon q^{n-1}$. For $\epsilon > 1/2$, this expression has a single root in the interval $(0, 1)$, given by

$$q^* = \left(\frac{2\epsilon - 1}{n\epsilon}\right)^{\frac{1}{n-1}}, \tag{S30}$$

which maximises $w_{\mathrm{R}}(q)$. Evaluating (S29) at $q = q^*$ and simplifying, we obtain that coordinated specialisation is favoured over random specialisation (i.e., $w_{\mathrm{C}}(p^*) > w_{\mathrm{R}}(q^*)$ holds) if $c_n < c^*$, where

$$c^* = 1 - \frac{w_{\mathrm{R}}(q^*)}{g_{n-1,n}} = 1 - \frac{1 - \epsilon + (2\epsilon - 1)\left(\frac{2\epsilon-1}{n\epsilon}\right)^{\frac{1}{n-1}} - \epsilon\left(\frac{2\epsilon-1}{n\epsilon}\right)^{\frac{n}{n-1}}}{1 - \epsilon + (2\epsilon - 1)(n-1)/n}. \tag{S31}$$

We present some numerical results in Fig. S3B, where we graphically show whether $c_n < c^*$ (i.e., coordinated specialisation is favoured) or $c_n > c^*$ (i.e., random specialisation favoured) holds, and where $c^*$ is calculated according to (S31). We find that that smaller group sizes (small $n$) and more essential cooperation (higher $\epsilon$) favour coordinated specialisation.

We also find that, when benefits depend on the number of helpers (Fig S3A) and group size is sufficiently small (small $n$), then less essential forms of cooperation (small $\epsilon$) may still favour fully coordinated specialisation. In contrast, when the benefits depend on the ratio of helpers to reproductives (Fig S3B) and the essentiality of cooperation is high (high $\epsilon$), then larger group sizes (large $n$) may still favour fully coordinated specialisation. Furthermore, when benefits depend on the number of helpers (Fig S3A) there seems to be little interaction between $\epsilon$ and $n$. That is, for any $\epsilon$ greater than a certain value (less than 1/2), the threshold value of $n$ seems to be the same. This is not the case when benefits depend on the ratio of helpers to reproductives (Fig S3B), where larger $\epsilon$ implies that a larger group size (higher $n$) is required to favour random specialisation.

# F Risk and insurance for random specialisers

Here, we show that more asymmetric fitness functions favour random specialisers to differentially weight their probability of adopting a helper role by making $q^*$ considerably different from $p^*$. This represents a form of insurance lessening the probability of forming particularly "risky" (low fitness) group compositions (Fig. S4A) [13, 14]. For such asymmetric fitness functions, the approximation that the optimal probability of adopting a helper role is equal to the optimal proportion of helpers is less accurate, and results using the approximation might be biased.

We use numerical simulations to generate random fitness functions without having to make many assumptions about their particular form. We generate 100,000 such fitness functions, some of which, by chance, will be more or less asymmetric than others. For a group size of $n$, we generate a fecundity sequence $f_{k,n} = (f_{0,n}, f_{1,n}, \ldots, f_{n,n})$ as follows. First, we sample $\epsilon$ from a uniform distribution (i.e., $\epsilon \sim U(0,1)$) and set the fecundity of a reproductive in a group with no helpers to $f_{0,n} = 1 - \epsilon$. Then, we draw $n$ numbers $h_1, \ldots, h_n$ from an uniform distribution (i.e., $h_i \sim U(0,1)$, for $i = 1, \ldots, n$), and set the fecundity of a group with $k > 0$ helpers to

$$f_{k,n} = f_{0,n} + \epsilon \frac{\sum_{i=1}^{k} h_i}{\sum_{i=1}^{n} h_i}. \tag{S32}$$

This creates a random fecundity sequence that monotonically increases from $1 - \epsilon$ to 1. Finally, we obtain the fitness sequence, $g_{k,n}$, by substituting S32 into S1 (Fig. S4B).

For each fitness sequence $g_{k,n}$, we numerically determine the optimal proportion of helpers, $p^*$ (S4), and the optimal probability of adopting a helper role for random specialisers, $q^*$ (S18). We reject any fitness function for which division of labour is not favoured (i.e., for which either $p^* = 0$ or $q^* = 0$ holds).

We measure the asymmetry of the fitness function as follows. We begin by setting

$$\underline{\gamma} = \sum_{k=0}^{k^*-1} \binom{n}{k} p^{*k} \left(1 - p^*\right)^{n-k} g_{k,n},$$

and

$$\overline{\gamma} = \sum_{k=k^*+1}^{n} \binom{n}{k} p^{*k} \left(1 - p^*\right)^{n-k} g_{k,n},$$

which respectively quantify the expected fitnesses of a random specialising group if it undershoots ($k < k^*$) or overshoots ($k > k^*$) the optimal proportion of helpers, under the assumption that the probability of becoming a helper is equal to the optimal proportion of helpers ($q = p^*$). We then use the ratio

$$\gamma = \frac{\underline{\gamma}}{\overline{\gamma} + \underline{\gamma}}.$$

as a measure of the asymmetry in the fitness function, with $\gamma > 1/2$ indicating fitness functions where more is to be gained by undershooting $p^*$ and with $\gamma < 1/2$ indicating the reverse. Accordingly, an asymmetry ratio of $\gamma = 1/2$ indicates that the fitness function is symmetric.

The results of this analysis are shown in Figs. S4C and S4D. We find that on average the more asymmetric a fitness function is, the more that random specialisers are favoured to underweight or overweight their probability of becoming a helper, that is, the larger the distance between $q^*$ and $p^*$, $q^* - p^*$ (Fig. S4C). In turn, the larger the optimal deviation in the probability of becoming a helper (the larger $|q^* - p^*|$), the more significantly random specialisers stand to gain fitness advantages by evolving to this optimised allocation of labour ($q = q^*$) (Fig. S4D).

This numerical exercise illustrates that in some cases, groups of random specialisers are strongly favoured to underweight or overweight the individual probability of becoming a helper, even though this means that they are less likely to produce groups with the optimal allocation of labour. This occurs so as to minimise the risk of producing groups with a particularly poor allocation of labour and so may be considered a type of insurance strategy akin to the fertility insurance observed in the biased sex ratio of some blood parasites [13, 14]. The model in the main text can produce asymmetric fitness functions. However, we find in section D that this phenomenon does not change the qualitative predictions of the model (Fig. S1C). That is, smaller group sizes (small $n$) and more essential cooperation (higher $\epsilon$) still favour fully coordinated specialisation.

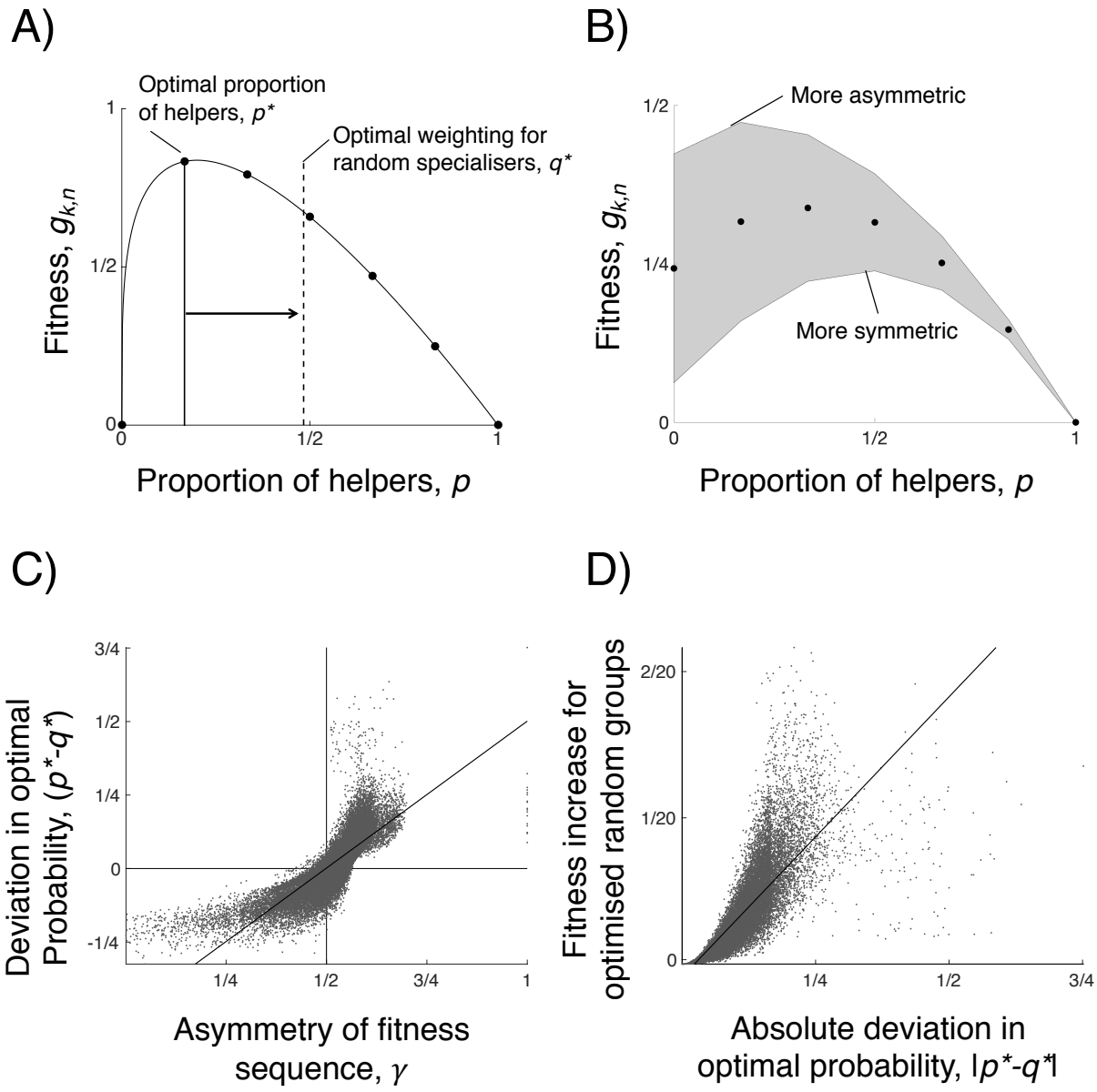


Figure S4: **Risk and insurance for random specialisers**. A) With asymmetric fitness functions, a group of random specialisers faces a different expected fitness cost from undershooting than from overshooting the optimal proportion of helpers, $p^*$. This can favour random specialisers to differentially weight the optimal probability of becoming a helper, $q^*$, as a form of insurance. Here we plot an arbitrary asymmetric fitness function, $g_{k,n} = (1 - k/n)k^{1/4}$, in which cooperation is essential. B) We randomly simulated 100,000 fitness sequences $g_{k,n}$ for a fixed group size of $n = 6$, some of which are more asymmetric than others. The 99% confidence interval over our simulations for each permissible proportion of helpers is shown in light grey. C) Across the 100,000 simulations, we found that the more asymmetric a fitness sequence is (i.e., the larger the $\gamma$ ratio) the more the optimal probability of adopting a helper role, $q^*$, is favoured to deviate from the optimal proportion of helpers, $p^*$. We plotted the best fit line through these data to illustrate the positive dependence ($\text{slope} = 1.0$, $\text{intercept} = -0.50$). D) Across the 100,000 simulations, we found that more significant departures in the optimal probability of adopting the helper role (larger $|q^* - p^*|$) led to a larger fitness increase for random groups that employ the optimal strategy ($q = q^*$). We plotted the best fit line through this data to illustrate the positive dependence ($\text{slope} = 0.13$, $\text{intercept} = -0.003$). This demonstrates selection for such an insurance strategy as a means to avoid "risky" group compositions.

# G The optimal level of coordination

In the first analysis of the main text, we assumed that all individuals in a coordinated group interact with one another and so have complete information about the phenotypes of their social partners when specialising. Consequently, coordinated groups are fully coordinated and always contain the optimal proportion of helpers to reproductives, $p^*$. Here we relax this assumption and solve for the optimal level of coordination in two separate analyses, based on different assumptions and using different methodologies.

## G.1 Individual based model

Here we model explicitly how increasing the number of cell-to-cell interactions within the group can lead incrementally to a more precise allocation of labour. We consider a costly trait $s \in [0, 1]$ (the level of coordination) that is equal to the independent probability that any two cells in the group interact with each other. When a focal individual specialises, it adopts a helper role depending on how the proportion of helpers amongst cells that it interacts with compares to a critical threshold, $t \in [0, 1]$. Variables $s$ and $t$ are co-evolving traits in our simulations; they influence both the degree to which groups are coordinated and the degree to which labour is divided (i.e., the realised proportion of helpers, $p$). As the expected connectivity of the group increases (i.e., higher $s$, more one-to-one interactions), there is a higher cost of coordination paid by the group. However, this higher cost of coordination may be offset by the increased amount of information afforded to each individual when specialising and thus an increased chance that the group ends up with the optimal proportion of helpers, $p^*$. This trade-off means that in some cases the optimal strength of coordination may in fact be intermediate ($0 < s < 1$), rather than the full coordination considered in the main analysis ($s = 1$).

In the following, we outline our individual-based model. First, we show how the simulations are initialised. Second, given a group with a particular level of coordination ($s$), we show how we determine which individuals interact with one another. Third, given a group with an interaction network and target proportion of helpers ($t$), we show how the allocation of labour within the group is determined. Fourth, given the allocation of labour, we show how the fitness of each group is determined. Finally, we describe how the individuals in the population compete globally from one generation to the next, and how mutation may affect the trait values of $s$ and $t$. Results for the evolved level of coordination and target proportion of helpers are depicted in Fig. 4 in the main text and in Fig. S6.

### G.1.1 Simulation initialisation

We create a population of approximately $N_\mathrm{T}$ cells, sorted into groups of $n$ cells. Specifically, we set the number of groups in the population as $N_\mathrm{g} = \lceil N_\mathrm{T}/n \rceil$, that is, the least integer larger than or equal to $N_\mathrm{T}/n$. Thus, the population size, $N_\mathrm{g} n$, is relatively constant while the number of groups increases as the group size decreases. This reduces the effect of demographic stochasticity across simulations. All individuals in the population are characterised by their trait values $s$ and $t$. We assume that all groups are founded by a single asexual individual, and thus that groups are clonal and that all individuals in the same group have the same trait values. At the beginning of the simulation, we assume that all individuals in the population have no propensity for either division of labour or coordination (i.e., $s = 0$, $t = 0$).

### G.1.2 Coordination network

Given a group with a particular level of coordination, $s$, we determine the between-cell coordination network by constructing a random graph $G(n, s)$ where the $n$ nodes correspond to the $n$ cells in the group and where each possible edge (representing the interaction between a given pair of cells) occurs independently with probability $s$ [15]. Thus, for a given $s$, the expected number of interacting pairs is $sn(n-1)/2$. If $s = 0$, there are no interacting pairs and individuals do not know the phenotypes of any other group members. If $s = 1$, all individuals interact with all other individuals in the group and know whether they are helpers or reproductives. In between these extremes, individuals may only interact with a subset of the cells in the group (or indeed none at all) and know only partial information about the proportion of helpers in the group.

### G.1.3 Individual specialisation

Before we can show how individuals within the group specialise in this framework, we need to establish a few definitions. For a given interaction network, $G$, let $V = \{v_1, \ldots, v_n\}$ be the set of nodes (individuals or cells) and $E$ be the set of edges, such that $e_{ij} \in E$ if individuals $v_i$ and $v_j$ interact. We define an *allocation of labour* as the map $a : V \to \{0, 1\}$ such that $a(v_i) = 1$ if individual $v_i$ is a helper and $a(v_i) = 0$ if $v_i$ is a reproductive.

The realised proportion of helpers to reproductives in the group is then given by $p = \sum_{i=1}^{n} a(v_i)/n$. We denote by $N_G(v_i)$ and $N_G[v_i]$ the open and closed neighbourhoods of individual $v_i$, respectively. The degree of $v_i$, $d_i$, is the number of cells in its open neighborhood, that is, $d_i = |N_G(v_i)|$. Using these definitions we can calculate the *open proportion of helpers in the neighbourhood* of individual $v_i$ as $p_i^{\mathrm{o}} = \sum_{v_j \in N(v_i)} a(v_j)/d_i$, that is, the fraction of its neighbours that are helpers. Similarly, the *closed proportion of helpers in the neighbourhood* of individual $v_i$ is $p_i^{\mathrm{c}} = \sum_{v_j \in N_G[v_i]} a(v_j)/(d_i + 1)$, that is, the fraction of its closed neighbours that are helpers. If an individual has no neighbours ($d_i = 0$), then the open proportion of helpers is not defined and the allocation of labour proceeds differently.

For a particular group (given a set of cells $V$ and an interaction network $E$), we determine the allocation of labour, $a$, using a threshold model. First, we assume that the group begins with an initial allocation of all reproductives ($a(v) = 0, \forall v \in V$). This implicitly presupposes that reproduction is the default phenotype of cells. Then for $\tau \geq n$ time steps, we randomly sample with replacement one cell $v_i \in V$ at a time from the group. We sample with replacement and set $\tau \geq n$ so that a cell that has "committed" to a particular phenotype may still switch if too many of its neighbours have adopted the same choice. This emphasises that the allocation of labour process is continuous in time and that developmental trajectories are plastic/responsive to their environment. Each chosen cell, $v_i$, considers the open proportion of cells in its neighbourhood $p_i^{\mathrm{o}}$ and chooses to adopt a helper or reproductive phenotype depending on its target proportion of helpers, $t$. If $p_i^{\mathrm{o}} < t$, then the cell adopts a helper phenotype ($a(v_i) = 1$) as there are fewer helpers in its neighbourhood than its target proportion of helpers. If $p_i^{\mathrm{o}} > t$, then the cell adopts a reproductive phenotype ($a(v_i) = 0$) as there are more helpers in its neighbourhood than its target proportion of helpers. If the open proportion of helpers is equal to the target proportion of helpers ($p_i^{\mathrm{o}} = t$), then the cell adopts either phenotype with equal probability. If the individual has no neighbours ($d_i = 0$), then the individual behaves as a random specialist with helper probability equal to the target proportion of helpers ($q = t$).

#### G.1.4 Fitness calculation

Once labour is allocated within the group, we calculate the fitness of the group using the fitness equation of the linear public goods model (S11), with the cost of coordination given by equation (S6), where the cost of coordination depends directly on the expected number of cell-to-cell interactions ($sn(n-1)/2$) rather than the realised number of cell-to-cell interactions because the former captures the effort that each individual puts into coordination instead of how successful its efforts actually were, and so is under evolutionary control. We consider alternate formulations for the cost of coordination in the subsequent subsection (G.2).

#### G.1.5 Creating a new generation

Each generation, we create the new population by sampling group founders from individuals of the previous generation. That is, we randomly pick individuals with a weighting equal to the relative fitness of each group. The process repeats $N_{\mathrm{g}}$ times with replacement. Each sampled individual forms a group of $n$ individuals with inherited trait values $s$ and $t$. When the founder individual is sampled, we assume that there is a chance $\mu$ of having a mutation in one of the traits. If a mutation happens, the trait value is perturbed by adding a normally-distributed random number, $\delta \sim \mathrm{Normal}(0, \sigma^2)$, where $\sigma^2$ is the variance in the size of the mutation, truncated between zero and one. Over time, the trait values of $s$ and $t$ across the population evolve towards an evolutionarily stable state.

#### G.1.6 Simulation results

We perform simulations for each combination of group size $n \in \{2, 4, \ldots, 38, 40\}$ and essentiality of cooperation $\epsilon \in \{1/2, 1/2 + 1/38, \ldots, 1/2 + 18/38, 1\}$. For each simulation we determine the number of groups as $N_{\mathrm{g}} = \lceil N_{\mathrm{T}}/n \rceil$, where $N_{\mathrm{T}} = 10,000$. All individuals are initialised with no coordination or no propensity for division of labour (i.e., $s = t = 0$). Next, we proceed in each generation by determining the interaction network of each group (given its strength of coordination, $s$), and then each group's allocation of allocation of labour (given its target proportion of helpers, $t$). We then calculate the fitness of each group and sample individuals for the next generation with a probability proportional to fitness. We assume that mutations occur with probability $\mu = 0.001$ and that the expected variance in the mutation size is $\sigma^2 = 0.1$. We set $\alpha = 0.1$ and $\beta = 0.01$ (parameters of the cost of coordination).

The results of the simulations are shown in Fig. 4D in the main text and in Fig. S5, where we examined 400 combinations of group size and essentiality and we ran the simulation 10 times for each parameter combination, and where each simulation lasted 30,000 generations. For each simulation, we denoted $\hat{s}$ and $\hat{t}$ as the average trait values across the population for a given generation. We denote $\bar{s}$ and $\bar{t}$ as the across-simulation average of $\hat{s}$ and

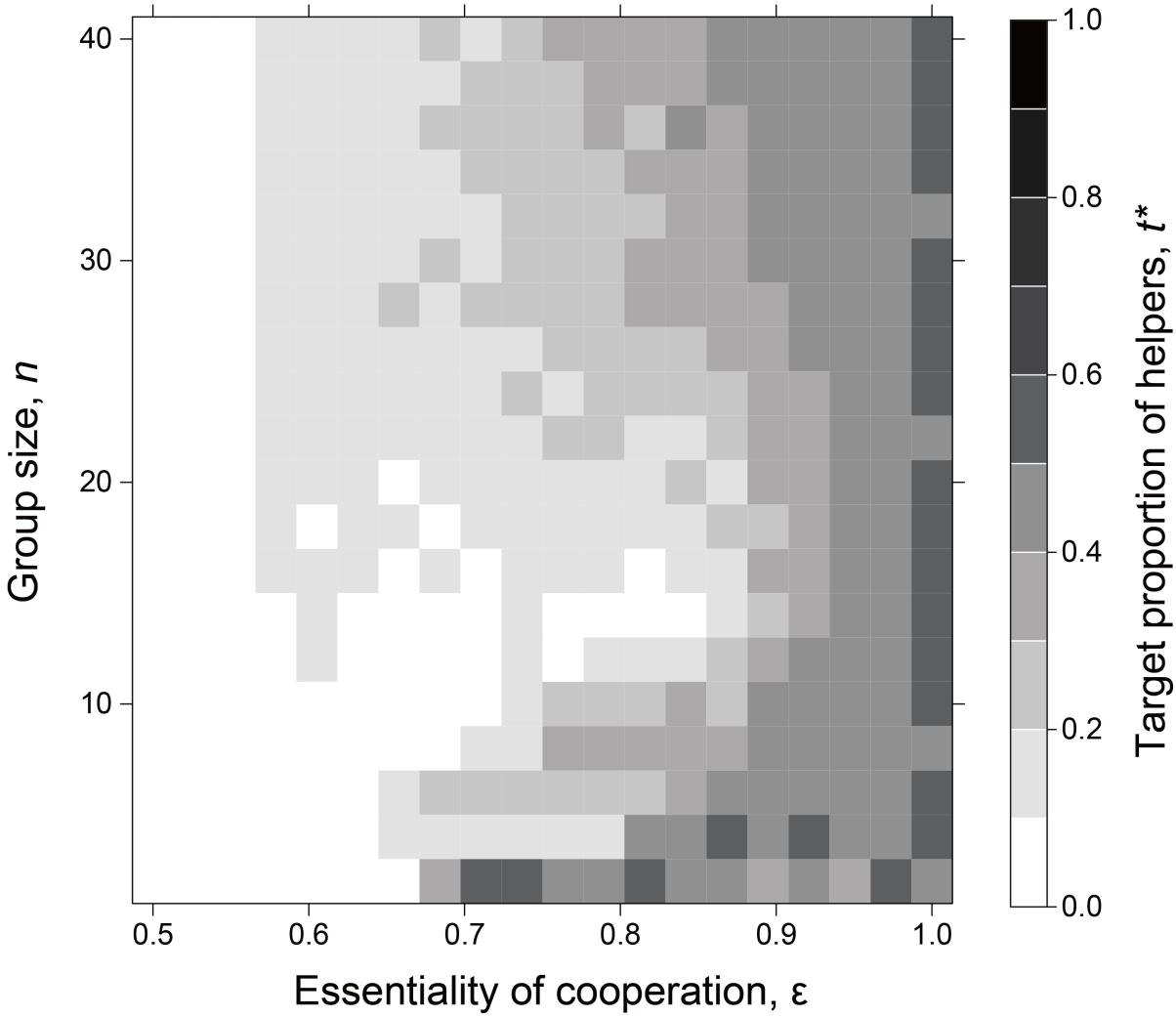


Figure S5: **Optimal target proportion of helpers**. We present the target proportion of helpers, $t^*$, that evolves in the same simulations used to produce Fig. 4D in the main text. We find that higher essentiality of cooperation (higher $\epsilon$) leads to a higher proportion of helpers but that group size ($n$) has little effect on the target proportion of helpers.

$\hat{t}$ for each generation (See Fig. 4B-C in main text). Finally, the evolutionary outcomes, $s^*$ and $t^*$, are estimated as the average $\bar{s}$ and $\bar{t}$ in the last 3,000 generations of the simulations. We show the evolutionary outcome $s^*$ as the main results (Fig. 4D in the main text) but also plot 10 individual time series of $\hat{s}$ and $\hat{t}$ and the across-simulation average $\bar{s}$ and $\bar{t}$ (Fig. 4B-C in the main text). To account for simulation stochasticity, we categorised random specialisation as any strategy for which $s^* < 0.1$, fully coordinated specialisation as any strategy for which $s^* > 0.9$ and intermediate specialisation as any strategy for which $0.1 < s^* < 0.9$. In Fig. S5, we show the results for $t^*$ in these simulations as well, finding that higher essentiality of cooperation (higher $\epsilon$) favours a higher proportion of helpers (higher $t^*$).

#### G.1.7 A note on our assumptions

There are several assumptions in this analysis that are worth expanding upon. First, we assumed that individuals specialise in a random order with replacement. This means that an individual may potentially change its phenotype several times during the course of the labour allocation. It may seem at first that this sort of phenotypic switching is not something that is observed in the natural world during the development of a social group. However, we do not envision here that individuals are actually differentiated until the end of the allocation of labour stage. Instead, while the allocation of labour is ongoing, the individuals are only committing to a developmental pathway and "signalling" to their partners which phenotype they "intend" to adopt. As the individuals in their neighbourhood do the same, it may be that an overabundance of one phenotype or the other is sufficiently strong to push the focal

individual into an alternate pathway. If we were to choose individuals in a random order without replacement, such that each individual only had one opportunity to adopt a helper role, this would introduce a systematic bias towards more helpers in the allocation of labour.

## G.2 Alternate costs of coordination

Throughout the presentation of our results, we have used the particular function for the cost of coordination given by (S6), with $\alpha = 0.1$ and $\beta = 0.01$. As it has been pointed out, the qualitative predictions (1-3) in the main text do not depend on the specific form that $c_n$ takes (so long as it is increasing in $n$). However, the optimal level of coordination that evolves might be sensitive to the form of the coordination cost (Fig. 4 in the main text). Here, we test the robustness of our results in Fig. 4 of the main text and section G.1 by repeating the simulations of section G.1 with alternate forms of coordination cost, $c_n$, to determine how the optimal level of coordination is affected.

We consider four functional forms for the costs of coordination $c_n$ that are each increasing in the expected number of cell-to-cell interactions $y = sn(n-1)/2$. Function I is a Type I functional response characterised by a linear increase in cost up to a maximum group size beyond wich the cost is equal to one, $c_n(y) = \min(1, \chi y)$. Second, we consider three functional forms characterised by an increasing but decelerating increase in cost, possibly up to maximum group size beyond which the cost is equal to one. Function IIa is the exponential form $c_n(y) = \alpha(1 - e^{\beta y})$, which we have employed in the analysis thus far (Fig. S6A). We set function IIb as $c_n(y) = \min(1, \psi(\xi y)^{\psi})$ (Fig. S6A). Finally, function IIc is a particular kind of type II response function: $c_n(y) = \min(1, \theta y/(1 + \phi y))$ (Fig. S6A).

For each functional form, we ran simulations to determine the optimal level of coordination, $s^*$. The results of these simulations are shown in Fig. S6B-E. Throughout we find that smaller groups (small $n$) and high essentiality of cooperation (high $\epsilon$) favour coordinated specialisation over random specialisation, with intermediate forms of specialisation ($0 < s^* < 1$) favoured to evolve at the interface between random specialisation ($s^* = 0$) and fully coordinated specialisation ($s^* = 1$). This is all in qualitative agreement with the findings of the main text.

We also find that the effect of the shape of coordination costs on the optimal mechanism depends on the functional form of the costs. When costs are set by function IIc, we found little effect on whether coordinated specialisation is favoured and the level of coordination that evolves (Fig. S6E), relative to linear costs. In constrast, when costs were determined by function IIa or IIb, we found both (i) a smaller likelihood of coordinated specialisation ($s^* > 0$) evolving and (ii) a relatively smaller chance that coordination will be intermediate ($0 < s^* < 1$), relative to full coordination ($s^* = 1$) (Fig. S6C and D).

The principal difference between function IIc and functions IIa and IIb is on the up-front costs of coordination. For both functions IIa and IIb, there is a large up-front cost to coordination before any coordination can evolve (Fig. S6A). In contrast, function IIc only has an initially linear cost of coordination (Fig. S6A). One consequence of a large up-front cost of coordination is that it is harder to evolve any coordination at all ($s^* > 0$). Another consequence is that intermediate coordination ($0 < s^* < 1$) may be less stable. Indeed, if coordination costs saturate quickly after a large up-front cost (as is particularly the case for function IIb, Fig. S6A), then it can be more advantageous for coordinating individuals ($s > 0$) to adopt full coordination ($s^* = 1$), having already paid a bulk of the costs associated with the mechanism. This can make the evolution of intermediate coordination comparatively rarer (Fig. S6C and D). The degree to which there are large up-front costs to coordination or how quickly costs saturate in the natural world will depend on the biology of specific species and the particular way that they might coordinate.

## G.3 Heuristic analytical model

In the main analysis (section C), we found that the fitness of a group of random specialisers with probability $q = p^*$ of producing helpers may be calculated as $w_{\mathrm{R}}(p^*) = G(p^*) - \epsilon \mathrm{Var}(P^*)$ (equation S13). That is, the fitness of a group of random specialisers is equal to the maximum possible fitness, $G(p^*)$, reduced by the hidden cost of random specialisation, which is proportional to the variance in the proportion of helpers across randomly specialising groups, $\mathrm{Var}(P^*) = p^*(1-p^*)/n$. Here, we sketch a simple, heuristic extension of this model with an evolving level of coordination $z \in [0, 1]$ that we assume has two opposing effects: (i) on the one hand, an increase in $z$ increases the total cost of coordination, and (ii) an increase in $z$ leads to a linear reduction in the hidden cost of random specialisation. To this end, we posit the fitness function

$$w(z) = (1 - c_n z)\left(G(p^*) - (1 - z)\epsilon \mathrm{Var}(P^*)\right), \tag{S33}$$

where $c_n$ is the proportional cost of increased coordination. Notice that $z = 0$ recovers the fitness function of the random specialisation mechanism, and $z = 1$ recovers the fitness function of the fully coordinated mechanism, where all groups have the optimal proportion of helpers, $p^*$. We deliberately do not specify how increasing

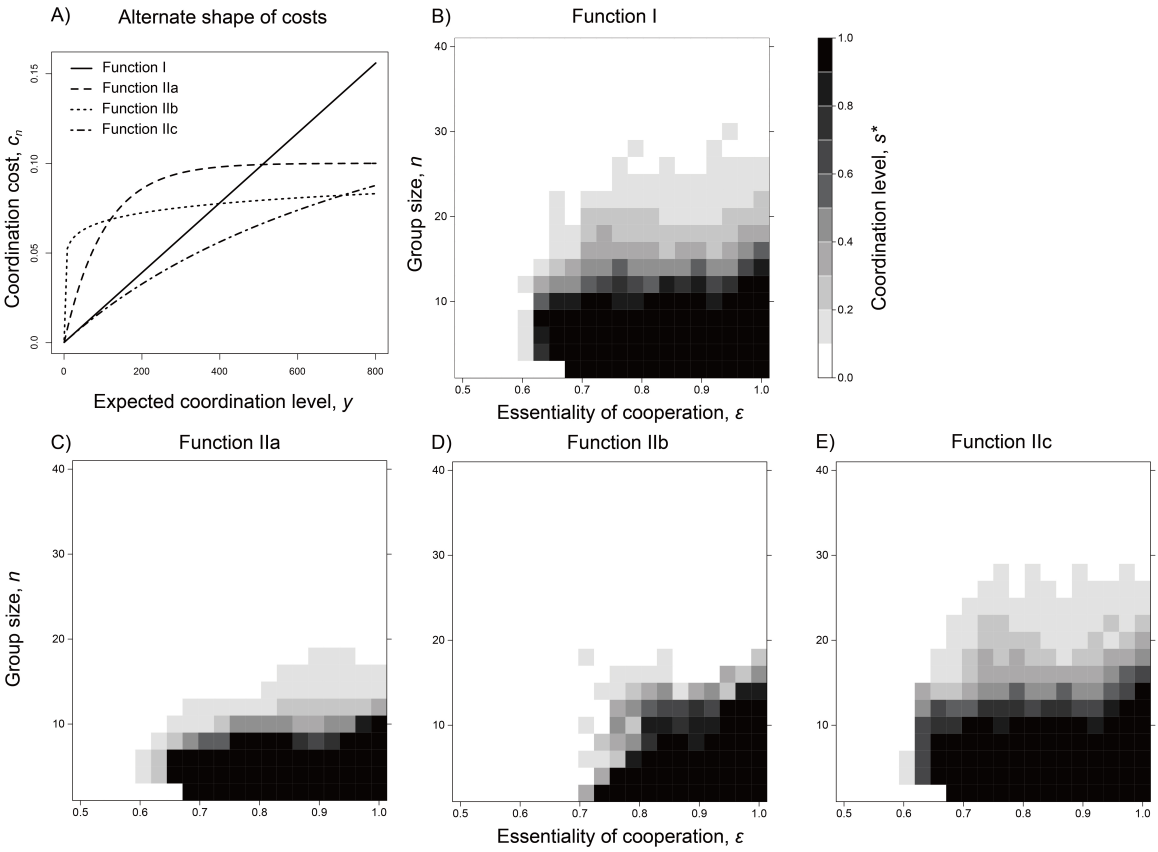


Figure S6: **The shape of the cost of coordination**. We show how the optimal level of coordination, $s^*$, depends on the form of the coordination costs, $c_n(y)$ as a function of $y = sn(n-1)/2$, the expected level of coordination. A) The four functional forms for the cost of coordination we explore. Function I is given by $c_n = \min(1, \chi y)$ ($\chi = 0.0002$); function IIa is given by $c_n = \alpha(1 - e^{\beta y})$ ($\alpha = 0.1$ and $\beta = 0.01$); function IIb is given by $c_n = \min(1, \psi(\xi y)^{\psi})$ ($\xi = 0.0002$ and $\psi = 0.1$); function IIc is given by $c_n = \min(1, \theta y/(1+\phi y))$ ($\theta = 0.0002$ and $\phi = 0.001$). B) The evolved level of coordination, $s^*$, when the cost of coordination is given by function I. C) The evolved level of coordination, $s^*$, when the cost of coordination follows function IIa. D) The evolved level of coordination, $s^*$, when the cost of coordination follows function IIb. E) The evolved level of coordination, $s^*$, when the cost of coordination follows function IIIc. Across all functional forms, we find that smaller groups (lower $n$) and more essential cooperation (higher $\epsilon$) favours coordinated specialisation. The degree to which intermediate forms of coordination ($0 < s^* < 1$) can evolve depends on the shape of the coordination cost.

$z$ decreases the hidden cost of random specialisation (i.e., the effective variance in the proportion of helpers), whether it be by increased cell-to-cell signalling, parental control or some other mechanism. The key feature of this heuristic model is the assumption of a simple tradeoff between producing potentially fitter but less precise groups on the one hand (lower $z$) and producing more precise groups at a larger coordination cost on the other hand (higher $z$).

Directional selection in $z$ is found by calculating the selection gradient $\mathcal{S}(z)$, which in this case is simply obtianed by taking the derivative of (S33) with respect to $z$:

$$\mathcal{S}(z) = w'(z) = -c_n G(p^*) + (1 + c_n)\epsilon \mathrm{Var}(P^*) - 2\epsilon c_n \mathrm{Var}(P^*) z \tag{S34}$$

Since $G(p^*)$ and $\mathrm{Var}(P^*)$ are independent of $z$, the selection gradient $\mathcal{S}(z)$ is linear and decreasing in $z$. The evolutionary dynamics are then fully specified by the sign of the selection gradient $\mathcal{S}(z)$ at the extreme points $z = 0$ and $z = 1$. To characterise the evolutionary dynamics, we start by writing conditions $\mathcal{S}(0) > 0$ and $\mathcal{S}(1) > 0$ respectively as $c_n < \underline{c}$ and $c_n < \overline{c}$, where the critical cost values $\underline{c}$ and $\overline{c}$ are given by

$$\underline{c} = \frac{\epsilon \mathrm{Var}(P^*)}{G(p^*) - \epsilon \mathrm{Var}(P^*)}, \tag{S35}$$

$$\overline{c} = \frac{\epsilon \mathrm{Var}(P^*)}{G(p^*) - \epsilon \mathrm{Var}(P^*)}. \tag{S36}$$

Using $G(p^*) = (1 - p^*)(1 - \epsilon + \epsilon p^*)$, $\mathrm{Var}(P^*) = p^*(1 - p^*)/n$, and approximating $p^*$ by $\hat{p} = 1 - 1/(2\epsilon)$ (see section C), $\underline{c}$ and $\overline{c}$ can be written as

$$\underline{c} = \frac{2\epsilon - 1}{n + 1 - 2\epsilon}, \tag{S37}$$

$$\overline{c} = \frac{2\epsilon - 1}{n - 1 + 2\epsilon}. \tag{S38}$$

For $\epsilon > 1/2$ these critical values satisfy $\underline{c} < \overline{c}$ and we arrive to the following characterisation of the evolutionary dynamics:

1. If $c_n \leq \underline{c}$ (low coordination costs), then $z = 1$ is the only convergence stable equilibrium, and full coordination evolves.

2. If $c_n \geq \overline{c}$ (high coordination costs), then $z = 0$ is the only convergence stable equilibrium, and no coordination evolves.

3. If $\underline{c} < c_n < \overline{c}$ (intermediate coordination costs), an intermediate level of coordination $z^* \in (0, 1)$ is the only convergence stable equilibrium. Solving $\mathcal{S}(z^*) = 0$, we find that the intermediate convergence stable level of coordination is given by

$$z^* = \frac{1}{2}\left(1 + \frac{1}{c_n} - \frac{n}{2\epsilon - 1}\right). \tag{S39}$$

Both critical costs are increasing functions of $\epsilon$ and decreasing functions of $n$. Likewise, $z^*$ is an increasing function of the essentiality of cooperation $\epsilon$, a decreasing function of the cost of coordination $c_n$, and, if $c_n$ is increasing with group size $n$ (which we assume), then $z^*$ is decreasing in $n$. It follows that lower costs of coordination (lower $c_n$), smaller groups (lower $n$), and more essential cooperation (higher $\epsilon$) all lead to (i) a higher likelihood that some coordination is favoured (i.e., that $c_n < \overline{c}$ holds), and (ii) a higher likelihood that full coordination is favoured (i.e., that $c_n \leq \underline{c}$ holds), and (iii) a higher evolved level of coordination (larger $z^*$) if intermediate coordination is favoured (i.e., $\underline{c} < c_n < \overline{c}$ holds). These results are in broad qualitative agreement with the results of the analysis in the main text where only perfect coordination and full random specialisation were considered, and with the individual-based model presented in section G.1.

# H Estimating the characteristics of social microbes

In Table 1 of the main text, we provided rough qualitative (verbal) estimates of relative group size, essentiality of cooperation, and the cost of coordination for labour dividing microbe species in which the mechanism of specialisation has been identified. This is not a quantitative, rigourous analysis. An across-species phylogeny-based comparative study with a larger species sampling would be required to formally test the predictions of our models. Here we provide further details on the reasoning behind some of our estimates in Table 1 of the main text.

We estimate relative group size by considering two key factors: how many individuals occupy a colony or aggregate of the species being considered, and how the benefits of cooperation are conferred to social partners. Fruiting bodies of *D. discoideum* and *M. xanthus*, populations of *B. subtilis*, *V. carteri* colonies, *S. enterica* co-infections and *A. cylindrica* filaments are potentially composed of hundreds to thousands of cells [5, 6, 7, 9, 10, 11, 12, 16, 17, 18, 19]. However, in some cases the benefits of cooperation are diffusive and so the effective radius of cooperation might be much smaller. For instance, protein degrading proteases diffuse locally in *B. subtilis* populations and fixed nitrogen must diffuse along the filament in *A. cylindrica* [5, 6]. Due to the geometry of the populations, this may lead to a medium sized effective social group in the former case (diffusion in multiple axes) and a small one in the latter (diffusion along one axis). In the other cases, the benefits of cooperation might be conferred directly to beneficiary cells, such as in *S. enterica*, *V. Carteri* and in the fruiting bodies, and so the effective size of the social group could be very large [7, 9, 10, 11, 12, 16, 17, 18, 19].

We approximate the relative essentiality of a trait by estimating whether reproductive cells would survive and proliferate in the absence of helpers. In *S. enterica* co-infections, the action of suicidal helpers only provides a competitive advantage and so might not be 'essential' [11, 12, 16]. In *B. subtilis* and *A. cylindrica*, the amount of degraded protein or fixed nitrogen in the environment is growth rate-limiting and so helpers might only provide an advantage to the replication rate of reproductive cells (non-essential cooperation) [5, 6]. Fruiting bodies of *M. xanthus*, and *D. discoideum* form when there are insufficient nutrients locally and so cells that do not partake in the complex cannot disperse to more favourable environments (more essential cooperation) [17, 18, 19]. Finally, *V. carteri* colonies that cannot stay afloat at the proper height in the water column may not receive enough light for photosynthesis, leading to cell death (essential cooperation) [9, 10].

We estimate the relative cost of coordination by considering two key factors: how close are cells to each other (attached or not attached), and how complex are the cells (how metabolically taxing would coordination be). For example, populations of *S. enterica* and *B. subtilis* are free-floating and so a signalling system would require a dispersed product [5, 11, 12]. Furthermore, they are both bacteria and so the relative metabolic costs could be quite high. In contrast, *M. xanthus* and *D. discoideum* form a multicellular slug before fruiting body formation and so coordination between cells would occur over a shorten distance [17, 18, 19]. Further, *D. discoideum* is a eukaryote and so the metabolic costs of signalling may be reasonably low relative to the higher costs of operating as a complex cell. Some special cases would lead to even lower costs of coordination. In *V. carteri*, offspring colonies form within the mother germ cell and so parental influence on cell fate means that no intercellular signalling system is necessary [9, 10]. In *A. cylindrica*, cells are attached to each other and the amount of fixed nitrogen in the environment serves as a cue for whether a cell should differentiate as a helper cell. Thus, the cost of coordination might only be due to the cost of detecting the amount of fixed nitrogen in the environment [6].

# References


[1] Eshel, I., Motro, U., and Sansone, E. (1997). Continuous stability and evolutionary convergence. Journal of Theoretical Biology *185*, 333–343.

[2] Christiansen, F.B. (1991). On conditions for evolutionary stability for a continuously varying character. The American Naturalist *138*, 37–50.

[3] Dionisio, F. and Gordo, I. (2006). The tragedy of the commons, the public goods dilemma, and the meaning of rivalry and excludability in evolutionary biology. Evolutionary Ecology Research *8*, 321–332.

[4] De Jaegher, K. (2019). Harsh environments: Multi-player cooperation with excludability and congestion. Journal of theoretical biology *460*, 18–36.

[5] Veening, J.W., Igoshin, O.A., Eijlander, R.T., Nijland, R., Hamoen, L.W., and Kuipers, O.P. (2008). Transient heterogeneity in extracellular protease production by Bacillus subtilis. Molecular systems biology *4*, 184.

[6] Rossetti, V., Schirrmeister, B.E., Bernasconi, M.V., and Bagheri, H.C. (2010). The evolutionary path to terminal differentiation and division of labor in cyanobacteria. Journal of Theoretical Biology *262*, 23–34.

[7] Flores, E. and Herrero, A. (2010). Compartmentalized function through cell differentiation in filamentous cyanobacteria. Nature Reviews Microbiology *8*, 39.

[8] Inglis, R.F., Hall, A.R., and Buckling, A. (2013). The role of 'soaking' in spiteful toxin production in Pseudomonas aeruginosa. Biology letters *9*, 20120569.

[9] Koufopanou, V. (1994). The evolution of soma in the Volvocales. The American Naturalist *143*, 907–931.

[10] Herron, M.D., Hackett, J.D., Aylward, F.O., and Michod, R.E. (2009). Triassic origin and early radiation of multicellular volvocine algae. Proceedings of the National Academy of Sciences *106*, 3254–3258.

[11] Ackermann, M., Stecher, B., Freed, N.E., Songhet, P., Hardt, W.D., and Doebeli, M. (2008). Self-destructive cooperation mediated by phenotypic noise. Nature *454*, 987.

[12] Diard, M., Garcia, V., Maier, L., Remus-Emsermann, M.N., Regoes, R.R., Ackermann, M., and Hardt, W.D. (2013). Stabilization of cooperative virulence by the expression of an avirulent phenotype. Nature *494*, 353.

[13] West, S.A., Smith, T.G., Nee, S., and Read, A.F. (2002). Fertility insurance and the sex ratios of malaria and related hemospororin blood parasites. Journal of Parasitology *88*, 258–264.

[14] Gardner, A., Reece, S., and West, S. (2003). Even more extreme fertility insurance and the sex ratios of protozoan blood parasites. Journal of Theoretical Biology *223*, 515–521.

[15] Gilbert, E.N. (1959). Random graphs. The Annals of Mathematical Statistics *30*, 1141–1144.

[16] Ackermann, M. (2015). A functional perspective on phenotypic heterogeneity in microorganisms. Nature Reviews Microbiology *13*, 497.

[17] Strmecki, L., Greene, D.M., and Pears, C.J. (2005). Developmental decisions in Dictyostelium discoideum. Developmental biology *284*, 25–36.

[18] Velicer, G.J., Kroos, L., and Lenski, R.E. (2000). Developmental cheating in the social bacterium Myxococcus xanthus. Nature *404*, 598.

[19] Strassmann, J.E., Zhu, Y., and Queller, D.C. (2000). Altruism and social cheating in the social amoeba Dictyostelium discoideum. Nature *408*, 965.