Biological and social systems are structured at multiple scales, and the incentives of individuals who interact in a group may diverge from the collective incentive of the group as a whole. Mechanisms to resolve this tension are responsible for profound transitions in evolutionary history, including the origin of cellular life, multicellular life, and even societies. Here, we synthesize a growing literature that extends evolutionary game theory to describe multilevel evolutionary dynamics, using nested birth–death processes and partial differential equations to model natural selection acting on competition within and among groups of individuals. We analyze how mechanisms known to promote cooperation within a single group—including assortment, reciprocity, and population structure—alter evolutionary outcomes in the presence of competition among groups. We find that population structures most conducive to cooperation in multiscale systems can differ from those most conducive within a single group. Likewise, for competitive interactions with a continuous range of strategies we find that among-group selection may fail to produce socially optimal outcomes, but it can nonetheless produce second-best solutions that balance individual incentives to defect with the collective incentives for cooperation. We conclude by describing the broad applicability of multiscale evolutionary models to problems ranging from the production of diffusible metabolites in microbes to the management of common-pool resources in human societies.

Life is intrinsically social at all scales. Even at cellular and subcellular scales, the incentives of a replicating individual may differ from those of their surrounding group, which presents a form of social dilemma. Evolutionary dynamics can operate at multiple scales of biological organization simultaneously, often creating tension between the fate of individuals and the groups in which they live. Mechanisms to relieve this tension are widespread and varied, and they have been the source of much complexity in biological phenotypes and behavior, including even the origin of cellular life itself (1–3). Failure to relieve this tension often results in pathology, disease, or social dysfunction (4–7).

Competition at multiple scales of organization is especially acute in biological systems on the cusp of major evolutionary transitions (8, 9), when the interests of individuals are at odds with the group incentives to form cooperative collectives (10, 11). This tension arises, for example, in the trade-off between an individual gene’s incentive for rapid replication versus the collective need to establish a balanced ensemble of genes for healthy cellular function—a conflict that can be mediated by genetic linkage, that is, by the origin of chromosomes (12–16), hypercycles (17), or package replication (18, 19) and stochastic correctors (20, 21). Multilevel competition continues to play a role in transitions to yet higher levels of biological organization, including the origins of multicellularity (22–24), such as the formation of multicellular colonies in Volvox (25) and slime molds (26, 27). In complex differentiated organisms, the breakdown of multicellular cooperation to suppress cellular-level cheating behaviors can manifest as disease, such as cancer, that threatens an organ or organism (4, 28–31). Multilevel competition is also found across vast scales of biological and social organization, in the evolution of so-called “evolutionary individuals” that constitute units of selection (32–34), from cells to societies.

Biologists have developed a robust theoretical framework to study strategic interactions in populations. Evolutionary game theory (35–39) originated from economic game theory, but it allows the study of dynamics and equilibria achieved by competition and differential reproduction, rather than by cognition and rationality. Biologists have uncovered a variety of mechanisms that can help reconcile individual incentives with behaviors that are beneficial to an entire group—such as assortative interactions, population structure, and reciprocity. These mechanisms can emerge by evolution among individuals possessing little or no cognitive capacity (36, 40–48). Even in a single group of
reproducing individuals, these mechanisms can promote cooperation or coordination to achieve collective outcomes that are far better than the Nash equilibrium of the underlying game. Nonetheless, these mechanisms may fail to promote any cooperation unless their impacts on individual fitnesses are sufficiently strong (47); and some mechanisms may produce a cooperative equilibrium only as a possible alternative to a stable, all-defector outcome (47, 49). These mechanisms that help promote cooperation under individual-level competition will interact with competition at multiple scales of biological organization, because they can impact both the individual advantage of defection and the collective benefit of cooperation (48, 50–53).

Evolutionary game theory can be generalized to describe natural selection operating simultaneously at multiple levels of organization in a group-structured population, characterizing both competition among individuals within a group and competition among groups of individuals (16, 53–68). This generalization provides a framework for exploring the countervailing effects of the individual incentive to defect and the collective incentive to cooperate, highlighting the evolutionary tug-of-war between levels of selection. Multilevel evolutionary game theory is relevant to questions as microscopic as the origins of chromosomes and cellular life, and as macroscopic as the dynamics of cultural evolution, cooperation, and conflict in human societies (16, 49, 61–66, 69–71).

In this paper, we synthesize recent developments in multiscale evolutionary game theory (53, 61, 62, 66, 67, 72); and we systematically contrast the outcomes of multiscale competition to those that can be achieved through mechanisms operating within a single group. We find qualitatively different long-term outcomes under competition in a single group, even with mechanisms such as assortment and reciprocity, than we find for simultaneous competition within and among groups. In most cases, however, neither within-group mechanisms nor among-group competition are sufficient to achieve collective optimality. And so, we also analyze the combined effects of within-group mechanisms such as assortment, reciprocity, and population structure with the effects of among-group competition. In most cases, these two mechanisms operate synergistically, providing greater collective benefits than expected under independence. We find that some within-group mechanisms, such as networked population structure, have qualitatively different effects on evolutionary dynamics when combined with among-group competition. And, in a few cases, the combination of within-group mechanisms and among-group competition can achieve collectively optimal outcomes.

Our presentation is structured as follows. First, we describe a general model of multilevel selection in evolutionary games, using a partial differential equation (PDE) to describe strategic compositions across many groups. We synthesize a body of results on the long-time behavior of this model, highlighting the strength of among-group competition required to support cooperation in steady-state; and we characterize the long-term average payoff achieved. Then, we incorporate assortment, reciprocity, and network structure within groups, and we determine the critical benefit-to-cost ratio required to sustain long-time cooperation when these within-group mechanisms operate in concert with multilevel selection. We then focus on multilevel selection when within-group interactions occur on regular social networks, showing that cooperation can sometimes be maximized when individuals have an intermediate number of network neighbors. Finally, we study a game in which payoffs depend on continuous levels of effort exerted by cooperators and defectors. We show how increasing cooperant effort increases both the collective incentive to cooperate and the individual incentive to defect, allowing us to find a “second-best” effort level that provides the best possible long-time collective payoff for a fixed strength of among-group selection. We conclude by discussing how our analysis of the tug-of-war between competition among individuals and among groups may be applied to study social dilemmas across scales, from genes to societies.

**Modeling Framework**

**Game-Theoretic Interactions.** To study the dynamics of multilevel selection, we generalize models from evolutionary game theory. We consider groups of individuals who engage in interactions with members of their peer group. Individuals can play one of two possible strategies, which we generically denote cooperate (C) or defect (D). In general, within a group containing a fraction \( x \) of cooperators, a cooperant has reproduction rate \( \pi_C(x) \) and a defector has reproduction rate \( \pi_D(x) \), for some arbitrary functions \( \pi_C \) and \( \pi_D \) satisfying \( \pi_D(x) > \pi_C(x) \).

We describe the evolution of behavior via individual-level selection in a large population by using the replicator equation, which tracks the proportion \( x \) of cooperators in the group over time, according to their relative rate of reproduction

\[
\frac{dx}{dt} = x (1-x) (\pi_C(x) - \pi_D(x)).
\]

This replicator equation can be derived either from models of reproductive competition via natural selection or from models of social learning based on individual-based decision-making (37, 38, 73, 74).

Much of our analysis holds for arbitrary frequency-dependent birth rates for two types, C and D. But to gain intuition about cooperation per se, we will often focus on pairwise games in which a cooperant pays a cost \( c > 0 \) to confer a benefit \( b > 0 \) to their opponent, whereas a defector pays no cost and confers no benefit. Assuming that two cooperants produce an additional payoff \( d \) representing possible (positive or negative) synergy for mutual cooperation (41), the outcomes of such a pairwise interaction are represented by the following payoff matrix

\[
\begin{pmatrix}
C & D \\
C & b - c + d & -c \\
D & b & 0
\end{pmatrix}.
\]

Payoffs to individuals are assumed to determine their rates of reproduction. When a group is composed of fractions \( x \) cooperators and \( 1-x \) defectors, individual payoffs to cooperators and defectors, averaged over all pairwise interactions, are given by \( \pi_C(x) = (b + d)x - c \) and \( \pi_D(x) = bx \), while the average payoff of group members is given by \( G(x) = (b -c)x + dx^2 \).

The payoff matrix of Eq. 2 corresponds to a prisoner’s dilemma (PD) if its entries have the following ranking

\[
b > b - c + d > 0 > -c
\]

(75). This ranking occurs provided that the synergy parameter \( d \) satisfy the condition \( c - b < d < c \). From the condition \( d < c \), we see that defection is a dominant strategy and that \( \pi_D(x) > \pi_C(x) \) for any \( x \) between 0 and 1. For PD games in which \( d < 0 \), the condition \( c - b < d \) tells us that \( b > c \), and
we can use this to show that the composition of cooperators \( x^* \) that maximizes average-payoff \( G(x) \) is given by

\[
x^* = \begin{cases} 
1 & : d \geq -\frac{b - c}{2} \\
\frac{b - c}{2d} & : d < -\frac{b - c}{2} 
\end{cases} \tag{3}
\]

Notably, intermediate levels of cooperation can maximize the collective outcome for the group when the synergy \( d \) of mutual cooperation is sufficiently negative.

The corresponding replicator equation within a group is

\[
\frac{dx}{dt} = x(1 - x)[\pi_C(x) - \pi_D(x)] = x(1 - x)(dx - c). \tag{4}
\]

For PD games (in which \( d < c \)), this replicator equation will always result in decreasing levels of cooperation, with the group converging to the all-defector composition \( x = 0 \) in the long-time limit. To study the dynamics of multilevel selection in evolutionary games, we will consider an analogue of the replicator equation that describes the effects of both within-group and among-group competition.

**A Nested Model of Multilevel Selection.** To study the evolution of strategic behavior at multiple levels of selection, we consider multiple groups, each internally engaged in the strategic evolution described above. In addition to within-group strategic change caused by the differential reproduction of individuals, we also imagine among-group dynamics that arise when one group decides to copy the strategic composition of a different group—or, equivalently, one group “dies” and another group “reproduces.”

Fig. 1 illustrates our model of multilevel strategic evolution, which follows the approach introduced by Luo et al. (61, 62, 76, 77). This figure provides an illustration of example individual-level and group-level birth–death events. One panel depicts the birth of a defector and the death of a cooperator within a single group, and the other panel depicts the birth of a two-cooperator group and the corresponding death of a three-defector group.

**A PDE Limit for Many, Large Groups.** From this individual-based model of simultaneous competition within and among groups, we can derive a continuum description of the dynamics in the limit of many of groups each of large size. The resulting PDE describes the evolution of the probability density \( f(t, x) \) of groups containing a fraction \( x \) cooperators at time \( t \):

\[
\begin{align*}
\frac{df(t, x)}{dt} & = -\frac{\partial}{\partial x} \left( (x(1 - x)[\pi_C(x) - \pi_D(x)]f(t, x) \right) \\
& \quad + \lambda f(t, x) \left[ G(x) - \int_0^1 G(y)f(t, y)dy \right]. \tag{5}
\end{align*}
\]

where \( \pi_C(x) \) and \( \pi_D(x) \) are the individual payoffs of cooperators and defectors and \( G(x) \) is the average payoff of group members in an \( x \)-cooperator group, and \( \int_0^1 G(y)f(t, y)dy \) is the average payoff of the whole population. The first term in this equation corresponds to within-group competition based on individual payoffs, and it pushes each group toward the all-defector composition in the prisoners’ dilemma (or toward a stable mix of cooperation and defection in other games, such as the Hawk–Dove game). The second term in Eq. 5 describes the impact of among-group competition, and it favors groups with high average payoff. The parameter \( \lambda \) governs the relative importance of within-group and among-group competition, namely the ratio of selection strengths for within-group and among-group competition. Eq. 5 is the multilevel analogue of the replicator equation presented in Eq. 4.

Eq. 5 is a first-order, nonlocal hyperbolic PDE, and so, we can use the method of characteristics (78, 79) to understand the multilevel evolutionary dynamics. The characteristic curves correspond to the traditional replicator equation for individual-level selection within each group, given by Eq. 4. And so, the characteristic curves describe the impact of within-group competition, while the solutions along characteristics track the effect of among-group competition. While the derivation of Eq. 5 is motivated by the payoffs generated by the game presented in Eq. 2, we can also use Eq. 5 to characterize the dynamics of multilevel selection for any continuously differentiable replication rates \( \pi_C(x), \pi_D(x) \), and \( G(x) \) for which defectors have an individual-level advantage over cooperators \( (\pi_D(x) > \pi_C(x)) \) and all-cooperator groups have a collective advantage over all-defector groups \( (G(1) > G(0)) \).
The method of characteristics provides the following representation formula for the solution $f(t,x)$ to Eq. 5 starting from an initial density $f(0,x) = f_0(x)$

$$f(t,x) = \exp\left(\lambda \int_0^t \left[ G(\phi_t^{-1}(x)) - \langle G(\cdot) \rangle f(x) \right] dx \right) \times f_0(\phi_t^{-1}(x)) \frac{d\phi_t^{-1}(x)}{dx}.$$  

[6]

Here $\langle G(\cdot) \rangle f(x)$ describes the characteristic curves solving the ODE

$$\frac{d\phi_t^{-1}(x)}{dt} = -\phi_t^{-1}(x) (1 - \phi_t^{-1}) \left[ \pi_C(\phi_t^{-1}) - \pi_D(\phi_t^{-1}) \right]$$

$$\phi_0^{-1}(x) = x,$$

which describes the effect of the within-group times solved backward in time from a given composition $x$ at time $t$. The derivation of the solution Eq. 6 to the multilevel dynamics is provided in SI Appendix, section 2.B. This formula shows that within-group dynamics push the initial distribution closer to the all-defector equilibrium, while among-group dynamics favor group compositions with high average payoffs. The long-time behavior of solutions $f(t,x)$ depends on the initial distribution of the strategic compositions across groups, $f_0(x)$, which we assume follows a uniformly continuous probability density. In particular, the dynamics depend on the density of the initial distribution near the full-cooperator equilibrium, $x = 1$, as quantified by the Hölder exponent $\theta$ at $x = 1$ (16, 62). A probability distribution with density function $u(t,x)$ has Hölder exponent $\theta$ near $x = 1$ provided

$$\lim_{y \to 0} \int_{-y}^y u(t,z) \, dz = \begin{cases} 0 & : \Theta < \theta \\ C & : \Theta = \theta \\ \infty & : \Theta > \theta \end{cases}$$

[8]

for some constant $C$. For example, the family of densities $f_0(x) = \theta (1 - x)^{\theta-1}$ have Hölder exponent $\theta$ near $x = 1$. Although the definition of the Hölder exponent $\theta$ is technical, we can think of $\Theta$ as the inverse of the size of the initial cohort of high-cooperator groups in the population.

The Long-Term Survival of Cooperation. Selection within a group tends to promote individuals who defect in the prisoner’s dilemma; whereas selection among groups tend to promote cooperative groups. Provided the among-group competition is sufficiently strong compared to within-group competition, that is $\lambda$ exceeds some threshold $\lambda^*$, then cooperation will survive over the long-term. This result, and all the results summarized in this section, hold not only for pairwise games (Eq. 2), but also for arbitrary frequency-dependent reproduction rates satisfying $\pi_D(x) > \pi_C(x)$ and $G(1) > G(0)$.

In particular, when $\lambda > \lambda^*$, the composition $f(t,x)$ converges to a steady-state density $f_0(x)$ that supports positive levels of cooperation, and that depends upon the initial density according to its Hölder exponent $\theta$. We prove this result in SI Appendix, Proposition 2.1, for point-wise convergence of density-valued solutions. Conversely, when among-group selection is too weak, $\lambda \leq \lambda^*$, all groups become pure defectors in the long-time limit (16)—that is, $f(t,x)$ converges to the delta-function $\delta(x)$ concentrated at $x = 0$. The threshold selection strength that governs the long-term survival of cooperation is given by the following formula*

$$\lambda^* := \frac{\left( \pi_D(1) - \pi_C(1) \right) \theta}{G(1) - G(0)}.$$  

[9]

We can interpret the numerator above as a defector’s payoff advantage in a group otherwise composed of cooperators, while the denominator describes the collective advantage of being a full-cooperator group rather than a full-defector group. And so, we see that the condition for cooperation to persist requires there be enough relative among-group selection to compensate for the temptation of an individual to defect in a group of cooperators. In addition, note that the threshold selection strength is increasing in the Hölder exponent $\theta$, meaning that a population with a smaller initial cohort of groups near full-cooperation requires stronger among-group competition to sustain any long-term cooperation.

We can also explore how multilevel competition improves collective outcomes by studying the average population payoff at steady state. When among-group selection is too weak, $\lambda \leq \lambda^*$, all groups converge to full defection, and the population average payoff is simply $G(0)$, the payoff of a full-defector group. But when group selection is sufficiently strong, $\lambda > \lambda^*$, the mean population payoff converges to $G(x)$ averaged over the steady-state density $f_0(x)$ given by\footnote{See (ref. 67, Eq. 55) and (ref. 72, Eq. 1.25).}

$$\langle G(\cdot) \rangle f_0 = \int_0^1 G(x) f_0(x) \, dx = G(1) - \frac{\theta}{\lambda} \left( \pi_D(1) - \pi_C(1) \right).$$

[10]

And so, in total, the long-term mean population payoff satisfies

$$\lim_{t \to \infty} \langle G(\cdot) \rangle f(t) = \begin{cases} G(0) & \text{when } \lambda \leq \lambda^* \\ G(1) - \frac{\theta}{\lambda} (\pi_D(1) - \pi_C(1)) & \text{when } \lambda > \lambda^*. \end{cases}$$

[11]

Combining this formula with Eq. 9, we can further decompose the long-term collective outcome as

$$\lim_{t \to \infty} \langle G(\cdot) \rangle f(t) = \begin{cases} \left[ \frac{x^*}{\lambda} \right] G(0) + \left[ 1 - \frac{x^*}{\lambda} \right] G(1) & \text{when } \lambda \leq \lambda^* \\ \frac{x^*}{\lambda} G(0) + \left( 1 - \frac{x^*}{\lambda} \right) G(1) & \text{when } \lambda > \lambda^*. \end{cases}$$

[12]

This shows us that the collective outcome interpolates between the all-defector payoff $G(0)$, when $\lambda \leq \lambda^*$, to the all-cooperator payoff $G(1)$, as $\lambda \to \infty$. In particular, this means that the population can never achieve a mean payoff greater than in a group composed entirely of cooperators, which occurs in the limit of infinitely strong among-group competition. For games in which the full-cooperator group maximizes group payoff, strong competition among groups can bring the population toward optimality. However, if a group’s payoff $G(x)$ is maximized by an intermediate proportion of cooperators, $x^* < 1$, then the population can never achieve an optimal payoff, even for arbitrarily strong among-group competition.

*See (ref. 67, Eq. 55) and (ref. 72, Eq. 1.25).

†See (ref. 67, Eq. 56) and (ref. 72, Eq. 2.8).
We call this phenomenon the shadow of lower-level selection: For some games, no amount of among-group competition can erase the fact that an individual defector sees an advantage in a group with many cooperators. This phenomenon also manifests in the distribution of cooperators achieved at steady state. In the limit of infinitely strong among-group competition, the long-time population will concentrate at a group composition featuring a level of cooperation that is less than the social optimum and that achieves the same collective payoff as the all-cooperator group. While the dynamics of Eq. 5 would produce a population concentrated at the level of cooperation that maximizes collective payoff in the absence of any within-group selection, we see that, in the presence of any within-group competition, the dynamics of multilevel selection feature a tug-of-war with forces acting at the equilibria of within-group dynamics. In other words, for any finite $\lambda$, within-group dynamics favor increasing levels of defection and a long-time collective outcome of $G(0)$, but the slow within-group dynamics near the all-cooperator composition allows the group-level success of all-cooperator groups with collective outcome $G(1)$ to promote cooperation. In SI Appendix, section 2, we provide a mathematical explanation for the long-time survival of cooperation via multilevel selection and why the long-time collective payoff of the population is limited by that of the all-cooperator group.

Although this PDE model for multilevel selection considers infinitely many groups of infinite size, the tradeoff between the collective incentive to cooperate and the individual incentive to defect also arises in finite populations. In SI Appendix, section 3, we consider a nested birth–death process introduced by Traulsen et al. (57–59) in a population with $m$ groups each composed of $n$ individuals, in which within-group competition acts on a faster timescale than among-group competition. For games with payoff matrix in Eq. 2, we find that fixation of cooperation is favored over fixation of defection if the ratio $W$ of selection strengths for group-level and individual-level replication exceeds the threshold

$$W^* = \frac{n-1}{2(m-1)} \left( (\pi_D(1) - \pi_C(1)) + (\pi_D(0) - \pi_C(0)) \right) \frac{1}{G(1) - G(0)}.$$  \(13\)

Similar to our PDE model, this threshold selection strength depends on a ratio between the individual-level and group-level payoff comparisons. The relative difficulty of promoting cooperation also depends on the ratio $\frac{n-1}{m-1}$ between the size of groups and number of groups, which plays a role in moderating or amplifying the tug-of-war between individual and group-level incentives. Due to the assumption of the separation of timescales between the individual-level and group competition in this stochastic model, some care is required in interpreting the relative selection strengths $\lambda^*$ and $W^*$ required for long-time success of cooperation in the PDE and stochastic models. We discuss these subtleties in SI Appendix, section 3.B.

### Simple Rules for the Evolution of Cooperation via Multilevel Selection

The results above show that the ability to support cooperation and achieve high collective payoffs via multilevel selection depends on the relative strength of within- versus among-group competition, $\lambda$, a defector’s advantage in an otherwise cooperative group, $\pi_D(1) - \pi_C(1)$, the success of full-cooperator groups $G(1)$, and the composition of cooperators that maximizes a group’s payoff, $x^*$. Because the ability to support cooperation via multilevel selection requires the collective incentive to cooperate to overcome the individual incentive to defect, we can explore biological mechanisms that promote cooperation either by improving the outcome of all-cooperator groups or by decreasing the individual-level temptation to defect among cooperators. We summarize seven such mechanisms in Table 1, and we devote the remainder of this section to studying how these mechanisms work in concert with multilevel selection to promote the evolution of cooperation.

The seven mechanisms outlined in Table 1 have already been studied in the context of a single level of selection (one group). Many of these mechanisms focus on ways to cluster cooperators with each other, or to incentivize cooperation in pairwise interactions. The first mechanism we consider is like-with-like assortment, in which individuals have a probability $r$ of interacting with individuals with the same strategy and probability $1 - r$ of facing a randomly chosen individual, as relatedness (42) or assortative matching (86) can help pair cooperators with other cooperators. Another mechanism that can promote cooperation is other-regarding preference, in which an individual’s utility or reproductive potential depends not just on their own payoff (with weight 1), but also on the payoff of their interaction partners (with weight $F < 1$) (36, 47, 48, 50, 80). We also consider mechanisms of reciprocity, in which cooperative individuals have a probability $q$ of detecting and punishing defectors based upon a reputation for defecting.

### Table 1. Summary of within-group mechanisms that may impact the tug-of-war between the collective incentive to cooperate and the individual incentive to cooperate

| Mechanisms          | Explanation                                                                 | References |
|---------------------|-----------------------------------------------------------------------------|------------|
| Assortment ($r$)    | Interact with same-strategy player with prob. $r$ and random individual with prob. $1 - r$ | (42)       |
| Other-regarding preference ($F$) | Place weight of $\frac{r}{1+r}$ own payoff and weight $\frac{F}{1+F}$ on payoff of opponent | (47, 48, 80) |
| Indirect reciprocity ($q$) | Cooperators identify defectors with probability $q$ and punish by defecting | (81)       |
| Direct reciprocity  | Individuals defect against those who have defected against them in prior interactions | (40, 82)   |
| Network reciprocity | Individuals play game and update strategy with local interactions on a $k$-regular graph | (46, 83)   |
| Continuous levels of effort | Payoff from public good depends on the level of effort characterizing cooperation and defection | (84, 85)   |

https://doi.org/10.1073/pnas.2216186120
We can be more precise in our analysis of multilevel competition by focusing on a prisoner’s dilemma with no synergy factor (direct reciprocity); or individuals have repeated interactions with the same partner with discount rate $\delta$ so that a player may punish defection by defecting in a future interaction (indirect reciprocity). We also analyze the effect of network structure for strategic interactions on $k$-regular graphs, that is, social networks in which each individual has $k$ neighbors. Finally, we consider a family of models in which payoff depends on the (continuous) level of effort expended by cooperators and defectors ($\theta$). In this model, contributing a higher level of effort increases the collective incentive to cooperate, but also introduces an increased incentive to defect against the more generous cooperators.

Each of these mechanisms, which have been previously studied in the context of a single level of selection, impacts the tug-of-war between individual-level selection for defectors and group-level selection for a high collective payoff. We summarize these mechanisms in Table 2, highlighting the effect of each mechanism on i) the individual-level incentive to defect in a group composed primarily of cooperators ($\pi_D(1) - \pi_C(1)$), ii) the strategic composition that maximizes the average payoff of a group ($x^*$), and iii) the collective incentive for a full-cooperator group over a full-defector group ($G(1) - G(0)$). Some of these mechanisms, such as assortment or other-regarding preference, decrease the individual-level incentive to defect in a cooperative group; whereas other mechanisms, such as birth–death dynamics on a network, increase this incentive; and other mechanisms, such as assortative matching have no effect on this incentive. Likewise, different mechanisms have different effects on the optimal level of cooperation within a group, as well as the collective benefit of pure cooperation.

Having summarized the qualitative effects of assortment and reciprocity in multilevel selection (Table 1), we now consider their quantitative effects within the broad class of three-parameter prisoners’ dilemma games (Eq. 2). From Eq. 9, the threshold for achieving steady-state cooperation in such games is given by

$$\lambda^{pD}_D = \frac{(e-d)\theta}{b-c+d} = \left(\frac{b}{b-c+d} - 1\right)\theta. \quad [14]$$

The threshold level of selection $\lambda^{pD}_D$ to support cooperation is a decreasing function of the synergy parameter $d$, so that increasing the synergy of mutual cooperation makes it easier to achieve cooperation via multilevel selection.

We can be more precise in our analysis of multilevel competition by focusing on a prisoner’s dilemma with no synergy factor ($d = 0$), which is called a simple donation game: cooperators pay a cost $e$ to confer a benefit $b$ to their opponent, while defectors pay no cost and confer no benefit:

$$C \begin{pmatrix} b - e & -c \\ b & 0 \end{pmatrix} D. \quad [15]$$

We reformulate the threshold condition on $\lambda$ (Eq. 9) by classifying games according to whether cooperation can or cannot survive under multilevel selection, given $\lambda$ and $\theta$. We characterize donation games by the minimum benefit-to-cost ratio $\frac{b}{e}$ for which cooperation is promoted by either individual or by multilevel selection. Table 3 summarizes simple rules for the preservation of cooperation, showing how the critical benefit-to-cost ratio can differ for multilevel versus single-level competition, and how it depends on the parameter $\frac{b}{e}$ of a multilevel process.

Table 3 shows how large a range of donation games support cooperation via multilevel selection, under each of the seven assortment and reciprocity mechanisms we have studied. For example, in the case of like-with-like assortment, the threshold benefit-to-cost ratio to support cooperation $\left(\frac{b}{e}\right)_{r,MS}^*$ is a decreasing function of the assortment probability $r$, because

$$\frac{\partial}{\partial r} \left(\frac{b}{e}\right)_{r,MS}^* = -\left(\frac{\lambda + \theta}{\lambda + \theta + r}\right)^2 < 0. \quad [16]$$

This means that, as the assortment probability increases, the combination of multilevel selection and assortment helps to promote cooperation across a broader range of donation games. And so, the presence of increasing assortment is always more conducive to cooperation under multilevel selection. By taking partial derivatives with respect to the other assortment and reciprocity parameters, we find that each such mechanism also has a monotonic effect—always decreasing the critical benefit-to-cost ratio for cooperation.

Table 3 also suggests a subtle aspect of the condition for persistent cooperation, even in the simplest case of well-mixed groups. Here cooperation is preserved provided $b/e > 1 + \theta$. This condition reduces to $b > e$ in the limit $\theta \rightarrow 0$; and the threshold $\lambda^{pD}_D$ required for cooperation approaches zero in this limit as well (Eq. 14). Intuitively, a probability distribution with Hölder exponent $\theta = 0$ near $x = 1$ behaves as if it contains a

| Table 2. Impact of mechanisms to promote cooperation via multilevel selection |

| Mechanisms                  | Effect on individual incentive to defect $\pi_D(1) - \pi_C(1)$ | Effect on optimal group composition $x^*$ | Effect on collective incentive to cooperate $G(1) - G(0)$ |
|-----------------------------|---------------------------------------------------------------|------------------------------------------|------------------------------------------------------|
| Assortment ($r$)            | $-$                                                          | $+$                                      | $0$                                                  |
| Other-regarding preference ($F$) | $-$                                                          | $0$                                      | $0$                                                  |
| Indirect reciprocity ($q$)  | $-$                                                          | $+$                                      | $0$                                                  |
| Direct reciprocity          | $+$                                                          | $+$                                      | $+$                                                  |
| $k$-regular graph death–birth (DB) updating | $+$ or $-$                                                  | $+$                                      | $0$                                                  |
| $k$-regular graph birth–death (BD) updating | $+$                                                        | $+$                                      | $0$                                                  |
| $k$-regular graph imitation (IM) updating | $+$ or $-$                                                  | $+$                                      | $0$                                                  |
| Level of defector effort ($e_D$) | $+$                                                          | N/A                                      | $-$                                                  |
| Level of defector effort ($e_C$) | $+$                                                          | N/A                                      | $-$                                                  |

The table summarizes the sign of the effect of each within-group mechanism on $\pi_C(1) - \pi_C(1)$, the individual-level incentive to defect; on $x^*$, the fraction of cooperators that maximizes group mean payoff; and on $G(1) - G(0)$, the collective incentive to achieve full-cooperation over full-defection. In the cases of $k$-regular graphs with death–birth or imitation updating, the sign of the effect on the individual incentive to defect, $\pi_C(1) - \pi_C(1)$, depends on the neighborhood size $k$ and the payoff matrix of the underlying game. For all other mechanisms considered, the impact on quantities that govern the outcome of multilevel selection have a single sign.
Table 3. Simple rules for the evolution of cooperation via multilevel selection

| Within-group mechanism | Threshold \( c \) (individual) | Threshold \( c \) (multilevel) | mechanism facilities multilevel cooperation |
|-------------------------|-------------------------------|-------------------------------|--------------------------------------------|
| Well-mixed (WM)         | \( \infty \)                  | 1 + \( \frac{\theta}{\lambda} \) | N/A                                        |
| Assortment (r)          | \( \frac{1}{r} \)              | 1 + \( \frac{(1 - r)\theta}{\lambda + \theta r} \) | Always                                     |
| Other-regarding preference (F) | \( \frac{1}{F} \) | 1 + \( \frac{(1 - F)\theta}{(1 + F)\lambda + F\theta} \) | Always                                     |
| Indirect reciprocity (q) | \( \frac{1}{q} \)              | 1 + \( \frac{(1 - q)\theta}{\lambda + \theta q} \) | Always                                     |
| Direct reciprocity (δ)  | \( \frac{1}{\delta} \)         | 1 + \( \frac{(1 - \delta)\theta}{(1 - \delta)\lambda + \delta \theta} \) | Always                                     |
| k-regular graph (DB)    | k                             | 1 + \( \frac{k(k - 1)(\lambda + k\theta)}{(k + 1)(k - 2)\lambda + k\theta} \) | \( k > \frac{2\lambda}{\theta} \) |
| k-regular graph (BD)    | \( \infty \)                  | 1 + \( \frac{k\theta}{(k - 2)\lambda} \) | Never                                      |
| k-regular graph (IM)    | \( k + 2 \)                   | 1 + \( \frac{(k + 1)k\theta}{(k + 3)(k - 2)\lambda + k\theta} \) | \( k > \frac{6\lambda}{\theta} \) |

The table summarizes the threshold benefit-to-cost ratios \( \frac{c}{\lambda} \) required to produce cooperation for a single level of competition (single group), or under multilevel selection, for either well-mixed within-group interactions, or for in-group dynamics that include a variety of different forms of assortment, reciprocity, or population structure. The last column indicates whether each mechanism of in-group structure facilitates cooperation across a broader class of games (lower \( b/c \) * in the multilevel setting, compared to well-mixed interactions in each group)—that is, whether the in-group mechanism is synergistic with multilevel competition in promoting cooperation.

delta-mass near \( x = 1 \)—that is, a positive portion of groups in the population is entirely composed of cooperators. In this special case, when the initial state is not even a uniformly continuous density function, previous work has shown that cooperation can eventually fix provided the all-cooperator group is favored over the all-defector group [(ref. 62, Theorem 3) and (ref. 72, Proposition 1)]. But this case falls outside the scope of analysis in this paper, which is focused on uniformly continuous densities (SI Appendix, section 2).

**Multilevel Selection with Network Reciprocity: Nonmonotonic Dependence on Neighborhood Size**

Compared to reciprocity and assortment, structured interactions along a network have a more nuanced impact on the evolution of cooperation in the multilevel setting. For a birth–death update rule in a structured population (46, 83) cooperation is never favored in the single-level setting (46, 83), regardless of the benefit-to-cost ratio, \( b/c \). But, cooperation can be supported in the multilevel setting for finite \( b/c > 1 + \frac{\theta}{(k - 2)^2} \). Nonetheless, this threshold is higher than the equivalent threshold for multilevel selection with well-mixed in-group interactions, \( b/c > 1 + \frac{\theta}{\pi} \). And so, graph structure with birth–death updating actually helps to inhibit the evolution of cooperation via multilevel selection.

Under either death–birth or imitation update rules (46, 83), the ability to support cooperation via multilevel selection depends on neighborhood size \( k \). Fig. 2 shows these effects in the case of death–birth updating (results are similar for imitation updating). For death–birth updating, the critical benefit-to-cost ratio required to achieve cooperation via multilevel selection exceeds the well-mixed threshold when neighborhood size \( k \) is sufficiently small (\( k < \frac{2\lambda}{\theta} \)), while the threshold is less than the well-mixed case when individuals have sufficiently many neighbors (\( k > \frac{2\lambda}{\theta} \)). By contrast, in a single group (single-level selection), cooperation is supported only if the benefit-to-cost ratio \( b/c \) exceeds the network degree \( k \) (46, 83), so the sparsest possible neighborhoods are most conducive to cooperation under individual-level selection. Consequently, placing interactions on a graph with sufficiently small neighborhood size \( k \) may help support cooperation at the individual level, yet hurt support for cooperation via multilevel selection. Furthermore, the critical benefit-to-cost ratio for multilevel selection with death-birth or imitation updating approaches that of the baseline (well-mixed) multilevel selection model in the limit of infinite neighborhood size (\( k \to \infty \)). As a consequence, the dependence of multilevel selection on node degree \( k \) is nonmonotonic: the space of donation games that can support cooperation via multilevel selection may be maximized by an intermediate neighborhood size \( k \) (Fig. 2, Bottom).

As these results show, the within-group dynamics on \( k \)-regular graphs are not simply endogenous versions of mechanisms of assortment and reciprocity, because they produce a variety of qualitatively different effects for cooperation via individual and multilevel selection.

In SI Appendix, section 4, we extend our analysis of multilevel selection with graph-structured in-group interactions to consider a broader class of prisoners’ dilemma games, characterized by payoff matrices (Eq. 2) with negative synergy \( d < 0 \). This class includes all PD games for which the group average payoff is maximized by a mix of cooperators and defectors. For birth–death updating, placing interactions on a \( k \)-regular graph always hinders cooperation via multilevel selection. For death–birth updating and imitation updating, however, multilevel selection with interactions on \( k \)-regular graphs can either facilitate or impede cooperation via multilevel selection. For a fixed neighborhood size \( k \), we characterize in SI Appendix, Propositions 4.2 and 4.4, the prisoners’ dilemma games with negative synergy for which within-group network structure increases or decreases the strength of among-group selection required for cooperation. We illustrate the parameter space of games for which these behaviors
can sometimes serve to promote full-cooperation in the limit of strong among-group competition (SI Appendix, Fig. S5); but that there are other games where the long-term level of cooperation will be lower than the collectively optimal value for any neighborhood size $k$ (the shadow of lower-level selection persists).

### A Model with Continuous Effort Levels

To explore the tug-of-war between collective and individual incentives in a setting with richer detail, we study a family of social dilemmas in which payoffs depend on a continuous levels of effort. The level of effort characterizes the degree of cooperation or defection (84). In this model, an individual exerting effort $e$ who interacts with an opponent exerting effort $\bar{e}$ receives payoff

$$u(e, \bar{e}) = a\bar{e} - e\bar{e}. \quad [17]$$

The socially optimal level of effort producing greatest collective payoff is $e^* = \frac{a}{2}$, although evolution by individual-level selection will never achieve this optimal outcome. Nonetheless, as we will see, multilevel selection can approach this social optimum.

We consider a pair of effort levels $e_C > e_D$ expended by a cooperator and defector, so that the utility function of a two-player two-action game (Eq. 17) is described by the payoff matrix

$$C \begin{pmatrix} e_C(a-e_C) & e_D(a-e_C) \\ e_C(a-e_D) & e_D(a-e_D) \end{pmatrix}. \quad [18]$$

In a single population, evolution will always lead to full defection, which produces the lowest possible payoff. In fact, if effort is a mutable and heritable trait then evolution in a single population will always decrease effort toward zero (87–89). Here, we study whether multilevel competition can support survival of the more cooperative effort level. A related question is whether the level of cooperation that multilevel selection supports ever produce the maximal group payoff, which is achieved when $e = \frac{a}{2}$.

Given the pairwise payoffs above, the collective payoff in a full-cooperator and full-defector group are given by $G(1) = e_C(a-e_C)$ and $G(0) = e_D(a-e_D)$, while the individual payoffs near the all-cooperator composition are given by $\pi_C(1) = e_C(a-e_C)$ and $\pi_D(1) = e_C(a-e_D)$. To guarantee that the all-cooperator outcome is favored over the all-defector outcome, we will assume that $G(1) > G(0)$, so the condition that $e_C(a-e_C) > e_D(a-e_D)$ (or, correspondingly, that $e_C + e_D < a$) results multilevel dynamics corresponding to a prisoner’s dilemma game. In this continuous-effort setting, then, the threshold selection strength $\lambda^*(e_C, e_D)$ required to sustain some long-time cooperation (i.e. some individuals contributing $e_C$) is given by

$$\lambda^*(e_C, e_D) = \frac{b\theta e_C}{a - e_C - e_D}. \quad [19]$$

Note that $\lambda^*(e_C, e_D)$ is increasing in $e_C$ and $e_D$, so that increasing the effort level of cooperators or defectors requires stronger among-group selection to support cooperation in the long-run.

The results above show that increasing the level of cooperation effort has two contrasting effects on cooperation. Increasing the cooperation effort $e_C$ makes sustaining cooperation more difficult, especially when among-group competition is weak, because it increases the individual-level incentive to defect $\pi_D(1) - \pi_C(1)$ more than it affects the collective incentive to cooperate $G(1) - G(0)$. But when the defector effort is larger, this increases...
the collective outcome $G(1)$ for the all-cooperator group, and thereby increases the best possible outcome achievable via multilevel selection—which is what matters when among-group competition is strong. We illustrate these two countervailing effects in Fig. 3, where for low $\lambda$, populations with lower cooperator effort achieve greater long-time collective payoffs (Fig. 3, Top), while populations with higher levels of cooperator effort achieve better collective outcomes when among-group competition is strong.

The question remains what level of cooperator effort $e_C$ maximizes the average payoff at steady state, for a given strength of among-group competition. Provided that $\lambda > \lambda^*$, the average payoff is given by (SI Appendix, section 5)

$$\langle G \rangle_f \sim e_C \left( a + \frac{\theta}{\lambda} \epsilon_D - \left( 1 + \frac{\theta}{\lambda} \right) e_C \right).$$

In the limit as $\lambda \to \infty$ the collective payoff of the steady-state population is $e_C (a - e_C)$, which is maximized for cooperator effort $e_C = \frac{a}{2}$. And so, for sufficiently strong among-group competition, multilevel competition not only supports some cooperation, but it actually approaches the socially optimal level of cooperative effort.

However, when $\lambda$ is finite, the socially optimal effort $e_C = \frac{a}{2}$ does not, in fact, maximize average payoff under the dynamics of multilevel selection. Instead, we find that there is an effort level $e_C < \frac{a}{2}$ that achieves the highest collective payoff under multilevel selection for given $\lambda$ (Fig. 3, Bottom). This level of cooperator effort can be seen as a second-best solution (90–92) that arises from the dynamics of multilevel selection. The point is that increasing cooperative effort increases both the collective incentive to cooperate $G(1) - G(0)$ and the individual incentive to defect $\pi_D(1 - \pi_C(1))$; and the second-best effort level $e_C$ provides the optimal balance of these opposing effects to promote collective payoff under multilevel competition.

In summary, our analysis shows that multilevel competition can support positive effort that would never evolve in a single group. However, even multilevel selection most favors an effort level that is second best to the socially optimal one, except in the limit of infinitely strong among-group competition.

**Discussion**

In this paper, we have synthesized results from a general model for multilevel evolution, and systematically compared their outcomes to those that arise under single-level selection. In the context of a social dilemmas, we have derived “simple rules” for how mechanisms of assortment and reciprocity, which are already known to facilitate cooperation within a single group, compare to and operate in combination with among-group competition. Our analysis highlights the key role of individual and average payoffs achieved at the all-cooperator and all-defector groups in determining the long-time support for cooperation.

Our results highlight how tradeoffs between individual and collective incentives shape the long-time prospects for cooperation under multilevel selection. In particular, we found that the mechanisms of assortment and reciprocity served only to decrease the individual-level incentive to defect, and had no impact on the average payoff of the all-cooperator group. This reduces the strength of among-group competition required to sustain long-time cooperation, but it does not allow for any improvement to collective payoff over the case of multilevel selection with well-mixed interactions when among-group competition is very strong. When within-group interactions take place on a $k$-regular graph, support for cooperation is maximized for graphs with an intermediate neighborhood size, in sharp contrast to the monotonic dependence on connectivity that is known to operate within a single group (46, 83). As this result demonstrates, the PDE replicator equation provides additional insights into the interplay of social dilemmas and population structure, revealing more nuance on the collective benefits of cooperation than can been seen by studying individual-level selection on its own.

We have also analyzed games in which payoffs to cooperators and defectors are parametrized by a continuous level of cooperator effort. In this case, unlike all others we have studied, there is a tradeoff between the collective incentive to cooperate and the individual incentive to defect, as we vary the effort expended by cooperators. For such games, we find that increased cooperator effort makes sustaining steady-state cooperation more difficult when among-group competition is weak, due to the increased individual-level advantage for defection; while at the same time the greater payoff to the all-cooperator group improves the
collective outcome achieved when there is strong competition among groups. In this setting, we can find second-best solutions for cooperative effort—such that an effort level that is socially suboptimal in a monomorphic population may still be the maximal collective payoff at steady state under multilevel selection. The key insight is that the “second-best” solution balances in the individual incentive to defect versus the collective incentive to cooperate, when cooperation levels are continuous.

The flexibility of our PDE model for multilevel selection provides room for future development and applications of this framework. While we have focused primarily on contexts in which payoffs arise from two-player, two-strategy symmetric matrix games, the method of analyzing the long-time behavior for our PDE carries through for any continuously differentiable individual-level and group-level replication rates that generalize the two-level evolutionary tension of the prisoners’ dilemma. And so, this framework can generate baseline predictions for the impact of competition within and among groups on the evolution of cooperative traits or behaviors under a range of biological and social mechanisms, including, e.g., public goods games.

Multilevel social dilemmas are widespread across biological and social systems, arising in settings such as the production of diffusible metabolic public goods (93–95), the supply of energy to a cell by mitochondrial DNA (96–98), replication control of plasmids (99), and protocol evolution (16). Such conflicts also arise in infectious diseases, where they may manifest in the evolution of virulence when individual-level incentives for a pathogen to replicate are misaligned with collective incentives to spread between host organisms (100–105). Explicit modeling of evolutionary dynamics operating at multiple levels of organization may be particularly relevant to empirical work on systems in which complementary genetic replicators compete at the individual level, such as in the case of heterotypic cooperation between viral strains (106) and genetic linkage of antibiotic resistance genes in bacteriophages (107), or the tension between defective interfering particles and full viral genomes for within-host and between-host viral spread (108–111). The models we have developed may also be relevant for studying cultural group selection, where simulation studies have been used to explore the evolution of altruistic punishment (52, 112), the evolution of social norms for indirect reciprocity (113), and the role of social mechanisms, including, e.g., public goods games.

There are many open questions about the behavior of our model for multilevel selection, motivated by realistic biological considerations that we have not yet addressed—such as mutation, migration, genetic drift, and demographic stochasticity. PDE models featuring some these phenomena have been introduced by other authors (68, 115–119). For example, Velleret found that a model with two-level selection and a Kimura diffusion term admits a unique quasistationary distribution. It is also possible to derive a system of ODEs describing from a two-level birth–death process, under the limit of infinitely groups each of finite size (16, 61, 66). Preliminary study of this ODE model reveals a unique quasistationary equilibrium, analogous to results on the diffusive PDE model of Velleret; and numerical simulations suggest that this quasistationary state can approach the density steady state of Eq. 5. When the initial population has H"older exponent $\theta = 1$ near $x = 1$. Exploring the connections between the steady-state behavior of our PDE model and the long-time behavior of stochastic and ODE models that account for finite-size effects is an important direction for future research.

Incorporating additional biological mechanisms will increase the difficulty of analyzing long-time dynamics in multilevel PDE models. Allowing migration between groups (115, 116), horizontal gene transfer (118), or horizontal transmission in a host-pathogen system (120, 121) will introduce a nonlocal nonlinearity into the advection term for individual-level dynamics. While some progress has been made on showing convergence to an all-defector state in a model with migration (122), these results assume weak competition among groups, and more work is required to understand the effects of migration in general. Another open problem is to extend our PDE model to include pairwise competition between groups, which produces more nonlinearity in the group-level replication term. Numerical simulations suggest that phenomena such as the shadow of lower-level selection persist in this version of the PDE model. But an important direction for future research is to mathematically analyze dynamics with pairwise among-group competition, to understand how nonlinear, nonlocal group-level replication influences cooperation. Future research may also adapt PDE models to explore games with asymmetric roles for players, repeated games with large action spaces, and between-group population structure (123).

Data, Materials, and Software Availability. There are no data underlying this work.

ACKNOWLEDGMENTS. We thank the anonymous referees for their constructive input. We acknowledge support from the Simons Foundation Math+X Grant to the University of Pennsylvania (J.B.P., Y.M., and D.B.C.), the John Templeton Foundation grant no. 62281 (J.B.P.), the National Science Foundation grants DMS-1514606 (D.B.C.) and DMS-2042144 (Y.M.), and the Army Research Office grant W911NF-18-1-0325 (D.B.C. and S.A.L.).
116. Y. Ogura, N. Shimakura, Stationary solutions and their stability for Kimura's diffusion model with intergroup selection II. *J. Math. Kyoto Univ.* **27**, 635–655 (1987).

117. J. F. Fontanari, M. Serva, Solvable model for template coexistence in protocells. *Europhys. Lett. (EPL)* **101**, 38006 (2013).

118. J. F. Fontanari, M. Serva, Effect of migration in a diffusion model for template coexistence in protocells. *Bull. Math. Biol.* **76**, 654–672 (2014).

119. J. F. Fontanari, M. Serva, Nonlinear group survival in Kimura’s model for the evolution of altruism. *Math. Biosci.* **249**, 18–26 (2014).

120. S. Van Vliet, M. Doebeli, The role of multilevel selection in host microbiome evolution. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20591–20597 (2019).

121. C. Pokalyuk, A. Wakolbinger, Maintenance of diversity in a hierarchical host-parasite model with balancing selection and reinfection. *Stoch. Process Appl.* **130**, 1119–1158 (2019).

122. D. B. Cooney. "PDE models of multilevel selection: The evolution of cooperation and the shadow of lower-level selection," Ph.D. thesis, Princeton University (2020).

123. A. Akdeniz, M. van Veelen, The cancellation effect at the group level. *Evolution* **74**, 1246–1254 (2020).