

# Collective action in the fraternal transitions

Jonathan Birch

Received: 29 September 2011 / Accepted: 27 January 2012 / Published online: 9 February 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** Inclusive fitness theory was not originally designed to explain the major transitions in evolution, but there is a growing consensus that it has the resources to do so. My aim in this paper is to highlight, in a constructive spirit, the puzzles and challenges that remain. I first consider the distinctive aspects of the cooperative interactions we see within the most complex social groups in nature: multicellular organisms and eusocial insect colonies. I then focus on one aspect in particular: the *extreme redundancy* these societies exhibit. I argue that extreme redundancy poses a distinctive explanatory puzzle for inclusive fitness theory, and I offer a potential solution which casts coercion as the key enabler. I suggest that the general moral to draw from the case is one of guarded optimism: while inclusive fitness is a powerful tool for understanding evolutionary transitions, it must be integrated within a broader framework that recognizes the distinctive problems such transitions present and the distinctive mechanisms by which these problems may be overcome.

**Keywords** Major transitions · Inclusive fitness · Kin selection · Eusociality · Multicellularity · Social evolution · Collective action · Redundancy

## Inclusive fitness and the major transitions

The power of inclusive fitness

The core insight of Hamilton's (1964, 1970) inclusive fitness theory is embodied in a famous remark by Haldane: "I would lay down my life for two brothers or eight

---

J. Birch (✉)  
Department of History and Philosophy of Science, University of Cambridge,  
Free School Lane, Cambridge, Cambridgeshire, UK  
e-mail: jgb37@cam.ac.uk

cousins". When interacting organisms share genes, they have two routes to genetic representation in the next generation: a direct route, through their personal fitness; and an indirect route, through the fitness of their social partners. By considering the indirect fitness effects of an individual's behaviour, weighted by its relatedness to the affected individual, we can explain cooperative phenomena that would otherwise defy evolutionary logic.

The power of inclusive fitness as a tool for explaining social evolution is aptly demonstrated by its ability to dissolve an evolutionary analogue of the Prisoner's Dilemma. In the standard, one-shot, two-player Prisoner's Dilemma, mutual cooperation is the only Pareto optimal outcome, but mutual defection is the only Nash equilibrium. On the face of it, any act that imposes a fitness cost ( $c$ ) on one player while conferring a benefit ( $b$ ) on its social partner will give rise to such a dilemma. To see this, consider the relevant payoff matrix:

	COOPERATE	DEFECT
COOPERATE	$b - c$	$-c$
DEFECT	$b$	0

Mutual cooperation is strongly Pareto optimal since, if either player were to switch from COOPERATE to DEFECT, its partner would be worse off. But DEFECT dominates the payoff matrix: an agent performs better by defecting than by cooperating, irrespective of the strategy its partner plays.

Things look bad for the evolution of cooperation. Crucially, however, we have neglected indirect effects. When we take these into account, weighting them by the relatedness ( $r$ ) between the players, the payoff matrix is transformed, and the prospects for cooperation seem considerably brighter (see Maynard Smith 1982; Nowak 2006; Taylor and Nowak 2007):

	COOPERATE	DEFECT
COOPERATE	$(b - c)(1 + r)$	$rb - c$
DEFECT	$b - rc$	0

In the new payoff matrix, mutual cooperation is still Pareto optimal, but mutual defection is not necessarily a Nash equilibrium. On the contrary, mutual cooperation will be the only equilibrium if and only if  $rb - c > 0$ . This is one form of 'Hamilton's rule' for the evolution of social behaviour.

Considering indirect fitness effects is not the only way to dissolve a one-shot, two-player Prisoner's Dilemma (see Nowak 2006; Taylor and Nowak 2007), but there are reasons to regard it as a particularly powerful approach. First, it allows for the evolution of behaviours that impose arbitrarily large costs ( $c$ ) on the actor performing them, provided the indirect benefits ( $rb$ ) are larger still. Hence, inclusive fitness can in principle explain *extreme altruism*, in which cooperators sacrifice all

(or virtually all) their personal fitness.<sup>1</sup> Second, it allows for the evolution of *non-reciprocal altruism*, since, if  $rb - c > 0$ , cooperation *dominates* defection: it is the superior strategy even if one's partner fails to express the altruistic gene. Both considerations point to the particular relevance of inclusive fitness to social evolution in the social insects, where extreme, non-reciprocal altruism among related individuals is rife (see, e.g., West-Eberhard 1975; Bourke and Franks 1995; Queller and Strassmann 1998).

### Old dog, new tricks

The power of inclusive fitness to explain the evolution of cooperation between related organisms is clear. In recent years, however, an emerging trend in evolutionary thought has presented this venerable theory with a new set of explanatory challenges.

The trend in question is a radical re-evaluation of the importance of sociality in the history of life. Cooperation, once thought to be confined to a relatively small number of animal taxa, is now considered an utterly central element of the Darwinian worldview (see, e.g., Queller 1997; Michod and Herron 2006; Calcott 2008; Queller and Strassmann 2009; Strassmann and Queller 2010; Bourke 2011). We can trace this shift to the booming interest in the 'major transitions in evolution', a research programme that began in earnest with the pioneering work of Maynard Smith and Szathmáry (1995).

Building on foundations laid by, among others, Buss (1987) and Bonner (1988), Maynard Smith and Szathmáry cast the history of life as a series of events in which integrated, higher-level individuals have evolved from collectives of lower-level entities. When we look at evolution in this new light, we start to see social phenomena where we saw none before: we see cooperation among cells within multicellular organisms, among organelles within cells, even among genes within a chromosome. As Andrew F. G. Bourke notes in his recent synthesis, *Principles of Social Evolution*:

Social evolution has grown outwards from the study of the beehive and the baboon troop to embrace the entire sweep of biological organization. It claims as its subject matter not just the evolution of social systems narrowly defined, but the evolution of all forms of stable biological grouping, from genomes and eukaryotic unicells to multicellular organisms, animal societies, and interspecific mutualisms (Bourke 2011, p. 7).

This dramatic expansion of the explanatory domain of social evolution theory naturally gives rise to a methodological question: will traditional approaches to social evolution extend straightforwardly to these new explanatory contexts? For instance, can theoretical tools designed to make sense of relatively simple social interactions *between* multicellular organisms turn out to explain the origins of the dazzling complexity *within* a multicellular organism?

<sup>1</sup> Here, and in the rest of the paper, I use 'cooperation' to denote any behaviour that confers an absolute fitness benefit on a social partner, and I reserve 'altruism' to denote any cooperative behaviour that also imposes an absolute fitness cost on the actor. This is in line with standard usage (see Hamilton 1964; Trivers 1985; Bourke and Franks 1995; West et al. 2007; Bourke 2011).

I want to explore these issues with particular reference to inclusive fitness theory. Inclusive fitness theory was not originally designed to explain the major transitions in evolution, but there is a growing consensus that it—perhaps more than any other approach in social evolution theory—has the resources to do so. We can again turn to Bourke for a very clear statement of this view:<sup>2</sup>

Hamilton's inclusive fitness theory (kin selection theory) provides a general theory of social evolution powerful and versatile enough to serve as the conceptual foundation for understanding the major transitions in evolution (Bourke 2011, p. 27).

Though Bourke champions inclusive fitness theory as a *general* theory of transitions in individuality, my focus in this paper will be more limited: I will concentrate in particular on the evolution of multicellularity, and of social complexity in the eusocial Hymenoptera. Following Queller (1997), I refer to these jointly as the *fraternal transitions*. In making this restriction, I am not assuming that inclusive fitness is irrelevant in the context of other transitions: on the contrary, provided we understand “relatedness” loosely enough, there is every reason to suspect that inclusive fitness will prove invaluable in explaining, for instance, the transition from lone replicators to replicators in compartments (see Michod 1983; Frank 1994; Szathmáry and Maynard Smith 1997). Rather, I focus on the fraternal transitions because they intuitively provide the ‘best case’ for the inclusive fitness paradigm, as they involve groups of closely related individuals. It seems reasonable to work on the assumption that the challenges which surface in the best case are likely to recur elsewhere.

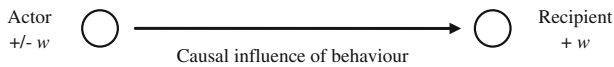
## Overview

In the next section, I discuss the qualitative differences between the cooperative interactions we find in the natural world's most complex societies—multicellular organisms and complex insect colonies—and those we see in simpler cases. In the two sections that follow, I highlight one feature in particular: the *extreme redundancy* exhibited by multicellular organisms and complex colonies. I argue that extreme redundancy poses a distinctive explanatory puzzle for inclusive fitness theory, and I offer a potential solution which casts coercion as the key enabler, at least in the case of the social insects. I suggest that the general moral to draw from the case is one of guarded optimism: while inclusive fitness is a powerful tool for understanding evolutionary transitions, it should be integrated within a broader framework that recognizes the distinctive problems such transitions present and the distinctive mechanisms by which these problems may be overcome.

## Proximate aspects of social complexity

The notion of inclusive fitness is usually introduced against a background of simplifying assumptions. We envisage a scenario in which one agent directly

<sup>2</sup> See also Maynard Smith and Szathmáry (1995); Queller (1997, 2000); Strassmann and Queller (2007); Queller and Strassmann (2009).



**Fig. 1** Cooperation as a pairwise fitness transaction

influences the fitness of another. The sign and magnitude of the fitness effects depend (relative to a given environment) only on the behaviour performed: the effects are not sensitive to what other agents in the population are doing. If agents interact multiple times, we can sum the individual effects to calculate the overall payoffs. In short, we assume a *pairwise fitness transaction* model of cooperation (Fig. 1). The two-player Prisoner's Dilemma is a familiar setup in which these assumptions hold.

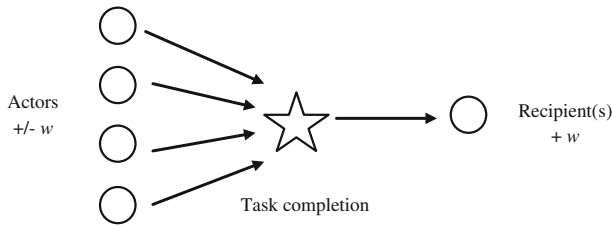
In fact, none of these assumptions needs to hold for the core idea of inclusive fitness—that is, the idea of considering indirect fitness effects as well as direct effects—to be usefully applied: inclusive fitness theory is not essentially wedded to a pairwise fitness transaction model of cooperation (see Gardner et al. 2011). Extending the theory to explain complex patterns of social interaction is not a straightforward business, however, and it is fair to say that our grasp of what the theory predicts in such cases remains limited.

In this section, I survey some of the distinctive ways in which the cooperation among cells in a multicellular organism, and among insects in the most complex eusocial colonies, departs from the pairwise fitness transaction model. These are features that distinguish social *complexity* from mere sociality, and which any theoretical framework that professes to explain transitions in individuality needs to be able to account for. In the subsequent sections, I explore the extent to which they can be explained by inclusive fitness considerations.

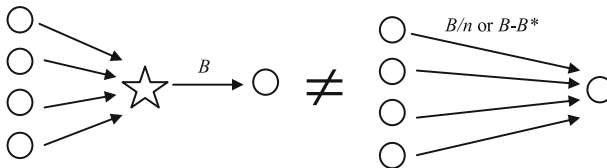
### Pairwise interactions versus collective tasks

To better understand the nature of social complexity, we can turn to the proximate literature on complex insect societies. The work of Carl Anderson, Nigel R. Franks and Daniel W. McShea is particularly valuable in this context (see Anderson and McShea 2001; Anderson and Franks 2001; Anderson et al. 2001). Anderson, Franks and McShea begin by observing that the most unmistakable feature of cooperation in complex societies—and the root of many further complexities—is that it takes the form not of isolated behaviours but of *collective tasks*, where a task is “an item of work that potentially [i.e., if completed] makes a positive contribution, however small, to inclusive fitness” (Anderson et al. 2001, 644).

While some tasks may be completed by a single individual, many require multiple contributions. The general picture in such cases is not the intuitive picture of a single actor conferring a benefit on a recipient: it is one of many actors collaborating to confer a benefit through task completion (Fig. 2). The recipients may be the same individuals as the actors (as in the case of a task performed by a number of individuals for their mutual benefit) but they need not be; indeed, in the kinds of cases that will concern us here—cases in which the recipient is a queen or a germ cell—the recipients rarely participate in any tasks, and the actors rarely gain any personal fitness benefit from their efforts.



**Fig. 2** Task-based cooperation



**Fig. 3** The benefit of task completion does not straightforwardly decompose into discrete, additive components contributed by the individual actors

Note that, even on the highly idealized picture of task-based cooperation shown in Fig. 2, it is unclear how, if at all, we can resolve the benefit conferred on the recipient into discrete components contributed by each of the actors. Should we say that each actor contributed  $B/n$ , where  $B$  is the total benefit conferred and  $n$  is the number of actors? Or should we say that each actor contributed  $B - B^*$ , where  $B^*$  is the reduced benefit that would have been conferred if that actor had not participated? The former measure takes no account of the fact that some actors may make a greater contribution to the task than others, while the latter measure allows that the total benefit conferred by the actors may differ from the total benefit received by the recipient.<sup>3</sup> Neither, therefore, is satisfactory. But if the total benefit of task completion cannot be resolved into discrete individual contributions, the pairwise fitness transaction model does not hold: task completion confers a benefit on the recipient that cannot be treated as a sum of the benefits conferred by the individual actors considered separately (Fig. 3).

Of course, the failure of these two simple measures hardly shows that the overall benefit *could not* be resolved into discrete components: some more complicated measure may yet succeed where the simple measures fail.<sup>4</sup> But it does show that

<sup>3</sup> Suppose, for instance, that every contribution is needed for the completion of the task, so that  $B^* = 0$  and  $B - B^* = B$ . On this measure, the benefit conferred is  $nB$ , but the benefit received is only  $B$ .

<sup>4</sup> Here is one possibility: each actor contributes a share weighted by the relative difference their contribution makes to the total benefit. Formally, let  $b_i$  represent the benefit conferred by the  $i$ th individual and let  $B_i^*$  represent the total benefit that would have been conferred had that individual unilaterally defected. The proposal is that:

$$b_i = \frac{B}{n} \cdot \frac{B - B_i^*}{\sum_i (B - B_i^*)}$$

This measure lacks the obvious defects of the simpler measures, but more work is needed to show that it provides a useful decomposition of the overall benefit for the purpose of understanding the relevant

decomposing the benefit of task completion is by no means a straightforward business: even in very simple cases, we can see how acknowledging the task structure of cooperation puts the pairwise fitness transaction model under strain.

### Further complications

I now want to introduce four further features that add to the complexity of task-based cooperation. While the features I will consider are by no means wholly absent from simpler forms of social group, complex societies tend to exhibit them to a much greater degree. The list is not intended to be exhaustive; moreover, since my primary aim is to relate these features to ultimate questions, I will describe each only briefly (see Anderson and McShea 2001 for a more detailed synthesis).

### *Division of labour*

Informal talk of division of labour is widespread in discussions of the major transitions (see, e.g., Maynard Smith and Szathmary 1995; Godfrey-Smith 2009; Bourke 2011), but Anderson, Franks and McShea deploy the notion in a relatively technical sense. For Anderson et al., labour is divided when a task is split into more than one distinct *subtask*, where a subtask is an item of work that would not by itself confer a inclusive fitness benefit but that fulfils one of the necessary conditions for the completion of a larger task. Subtasks may themselves be divided into further subtasks, and so on. Anderson, Franks and McShea provide no algorithm for the individuation of subtasks, but suggest that in practice the subtasks are often easily identified. For instance, they describe a grass harvesting task in *Hodotermes mossambicus*, where the workforce is visibly divided into cutters and transporters (Anderson et al. 2001, p. 645).

By conceptualizing division of labour in this way, we make the notion distinct from that of *specialization* (see below). Indeed, they are properties of different things: *tasks* are divided, while *workers* are specialized. This conceptual distinction, though rarely drawn explicitly, is a helpful one, because the division of a task into subtasks may occur without the specialization of workers, and vice versa.

### *Specialization*

Specialization is correlation between the properties of workers and the tasks they undertake.<sup>5</sup> It thus requires some form of *differentiation* among workers. In the eusocial Hymenoptera, two kinds of specialization predominate: specialization based on *morphological* differences (in which workers undertake different tasks depending on their physical characteristics) and specialization based on *age*

---

Footnote 4 continued

evolutionary dynamics. For instance, does Hamilton's rule still apply when benefit is calculated using this measure? I will not undertake this work here, since it is peripheral to the overall argument.

<sup>5</sup> Because specialization may be regarded as a kind of correlation, we can quantify the overall degree of specialization in a social group using information theory (see Gorelick et al. 2004). It is thus perhaps the only aspect of social complexity for which a reasonably straightforward quantitative measure is available.

differences (in which workers typically perform different tasks at different life stages). As Anderson and McShea (2001) note, however, some cases of specialization fall in between these categories. These are cases in which workers develop traits that enable them to perform a particular task at a particular life stage, only to lose those traits subsequently (they cite the short-lived production of royal jelly in honey bees, which leads to the temporary specialization of workers in feeding tasks). Polymorphism and age-based polyethism can thus blur into one another: we can do better by seeing these as extremes of a continuum of specialization based on developmental differences, ranging from superficial, short-lived differences in the simplest colonies to spectacular, life-long polymorphism in the most complex.

I want to reserve the term *extreme* specialization for cases in which specialization is accompanied by a loss of behavioural totipotency—in other words, cases in which workers have lost the ability to undertake some or all tasks other than the task for which they are specialized. While eusocial societies with distinct morphological castes exhibit some degree of extreme specialization, multicellular organisms display this phenomenon to a far greater degree: consider, for example, a human red blood cell, which specializes so exclusively in oxygen transport that it lacks even a nucleus, a basic prerequisite for participation in most other tasks.

*Germ-soma* specialization occurs when some group members specialize in tasks which contribute to the growth and persistence of the collective, while others specialize in tasks which generate new collectives. Germ-soma specialization may be extreme, such that somatic specialists lose the capacity to generate new collectives, but it need not be (in plants, for instance, all cells in the floral meristem can potentially give rise to new individuals; see Clarke 2011). Owing to its consequences for within-group conflict, germ-soma specialization is often assigned special importance in accounts of the major transitions (see, e.g., Buss 1987; Michod 2007; Godfrey-Smith 2009; Bourke 2011).

### Coordination

Coordination is a feat of signalling and plasticity, and introduces yet more contingencies on which the success of a task may depend: when a task requires coordination, the subtasks must be performed at the right time and in the right order. As Anderson and Franks (2001) take pains to point out, while coordination presupposes a division of labour, it may not always require specialization: a task must be split into subtasks, but the workers who undertake the subtasks need not belong to different specialized castes.

Among coordinated tasks, Anderson and Franks distinguish *partitioned* tasks, in which the subtasks take place in a coordinated *series*, from *team* tasks, in which the coordinated subtasks occur *concurrently*. While partitioned tasks are fairly widespread in eusocial societies (particularly tasks which exhibit a “bucket brigade” style organization; see Ratnieks and Anderson 1999), team tasks appear to be relatively rare. Anderson and Franks cite nest construction in *Oecophylla* weaver ants, prey retrieval in *Eciton burchelli* and *Dorylus wilverthi* army ants, and the decapitation of intruders in *Pheidole pallidula* ants. They are far from rare in



multicellular organisms, however, where teamwork is rife: consider a coordinated muscle contraction, or the coordinated production of enzymes.

### *Redundancy*

A workforce contains *redundancy* when there are more workers than are strictly needed for task completion. We see two broad kinds of redundancy in insect societies. The first sort (which I will call *passive redundancy*) occurs when there is a large reserve workforce, idle but ready to step in should any labour shortages arise. This phenomenon is widespread in eusocial societies (see Hölldobler and Wilson 1990, pp 342–343). The second (which I will call *active redundancy*) occurs when more workers actively undertake a task than are strictly necessary for its completion. We see this in the foraging strategies of complex ant societies: large numbers of ants search for food in parallel, then work in parallel to retrieve the food that one individual has found (see Oster and Wilson 1978, Herbers 1981). The upshot of redundancy in either form is that “if one worker doesn’t complete the task someone else will” (Oster and Wilson 1978; see also “[Why redundancy matters](#)”).

We see a clear analogue of this phenomenon in multicellular organisms, where the number of cells that specialize in a given task often dramatically exceeds the minimum required for task completion. To take a particularly extreme example, the human circulatory system can stand to lose one eighth of its total stock of red blood cells during a routine blood donation without any significant adverse effects.

Redundancy will receive further attention below. In the present context, I want to distinguish what I have called redundancy (following Anderson and McShea 2001, and Hölldobler and Wilson 1990) from a very different phenomenon to which the same name has been applied. Bourke and Franks (1995, p. 440) contrast what they term the *redundancy of parts* (that is, the existence of surplus workers, which I am calling simply *redundancy*) with what they term the *redundancy of functions*. By “redundancy of functions”, they mean an individual worker’s latent capacity to undertake tasks that they are never called upon to perform during their lifetime. I will call this phenomenon *latent versatility*. The issue is terminological but important nonetheless, since the contrast between redundancy and latent versatility will play a key role in “[Why redundancy matters](#)”.

From proximate to ultimate

Even a very brief overview of the complexities of cooperation exposes the limitations of the pairwise fitness transaction model. Many of the most striking features of social complexity depart from this model in significant respects. Of course, this is not to say that traditional approaches to the evolution of cooperation are of no use in the context of the transitions. It is merely an acknowledgement that, when we switch to a transitions context, new explananda come into view. An adequate theory of evolutionary transitions in individuality must be able to explain the complexities of task-based cooperation. It is, at present, an open question whether inclusive fitness theory has the resources to meet that challenge. To get

closer to an answer, I want to home in on one particular aspect of social complexity that has gone largely neglected in previous discussions of the major transitions.

### Why redundancy matters

In the opening pages of their seminal work, *Caste and Ecology in the Social Insects* (1978), Oster and Wilson identify redundancy (along with caste specialization and division of labour) as a fundamental feature of complex sociality (see Oster and Wilson 1978, pp. 11–15; see also Hölldobler and Wilson 1990, pp. 356–358). As Oster and Wilson emphasize, redundancy (or, in their terms, task structures based around parallel operations) is essential to the stability of a colony, and instrumental in generating benefit for social over solitary living.

The reason is that redundancy assures the *robustness* of task completion in the face of individual failure. When overall task success requires that *every* worker completes its subtask, the overall reliability (that is, the probability of task completion) can be no greater than the product of reliabilities of the individual workers. By contrast, when the task structure contains redundancy (whether active, passive, or both), the overall reliability can vastly exceed this product. When duplicate roles are performed in parallel, or when reserves are ready to step in should individual workers fail, overall task completion may remain highly probable even given a reasonably severe rate of worker attrition.

Redundancy, then, generates benefit for group living by enhancing the robustness of task completion.<sup>6</sup> But it may have an even more significant role to play in evolutionary transitions: the achievement of robustness through redundancy—rather than through latent versatility—may help facilitate the evolution of extreme specialization.

To see why, imagine a society in which there is extreme caste-based specialization such that (1) a member of one caste cannot perform the tasks for which other castes are specialized, and (2) each caste has exactly the minimum number of members required for the successful completion of its own specialized task. Such an arrangement would be highly prone to task failure in the face of chance events: if a single worker were killed or incapacitated, that worker's caste would no longer be able to function effectively. To secure the robustness of task completion in the face of individual failure, the workforce has two broad options. One is *latent versatility*: if workers of other castes are able to step into replace their incapacitated colleagues when not required for their own specialized task, task completion will not be imperilled by the failure of a single worker. The other is *redundancy*: if roles are duplicated beyond necessity, or if there are reserve workers on standby, the workforce will again be buffered against individual failure. If a group secures robustness primarily through this second route, the group will be able to sacrifice latent versatility for the sake of more efficient specialization without thereby becoming vulnerable in the face of environmental vicissitudes.

<sup>6</sup> See Calcott (2008) for a discussion of various ways in which cooperation may generate benefit for group living over social living. One of these ways, “reducing risk”, can be viewed as a form of redundancy; see below.

In this sense, redundancy may act as an enabling condition for the evolution of extreme specialization. This conjecture is supported by the correlation between redundancy, specialization and the loss of behavioural totipotency that we see in the eusocial Hymenoptera (Anderson and McShea 2001). Even the largest and most complex colonies, however, retain a large generalist caste of considerable latent versatility (Mirenda and Vinson 1981). For a more extreme data point we can look to paradigm multicellular organisms, which exhibit unparalleled levels of both specialization and redundancy: consider again the example of a red blood cell. In both cases, it seems reasonable to hypothesize that the transition from behavioural totipotency to extreme specialization was accompanied by a transition from robustness-via-versatility to robustness-via-redundancy.

Redundancy is not the sole preserve of eusocial colonies and multicellular organisms. It is also present to some degree in much simpler societies: for instance, animals from the same social group may each hunt independently and share the spoils among their fellow group members (a phenomenon observed in vampire bats; see Anderson and Franks 2001 and Calcott 2008 for discussion of such cases).<sup>7</sup> This can be viewed as a very simple form of redundant task structure: the task is to provide sufficient food for the group, and a number of individuals undertake the task in parallel. The benefit for the group, as before, is robustness in the face of individual failure. Complex societies, however, tend to exhibit redundancy to a far greater degree and achieve a greater degree of robustness as a result; indeed, the degree of redundancy is often so extreme that an individual contribution taken in isolation makes no significant difference at all to the probability of task completion. The difference between simple and extreme redundancy is one of degree rather than kind, and as such is compatible with a gradual transition from simple to complex social organization. As I will argue presently, however, this does not make the evolution of extreme redundancy easy to explain.

## Redundancy as a puzzle

### Collective action problems and the ‘paradox of voting’

In classical game theory, the phrase ‘collective action’ is almost invariably followed by ‘problem’. Suppose there is some task, the completion of which represents a Pareto optimal outcome. If the focal agent participates in the task, she incurs a cost. But if the task is completed, she receives the same benefit regardless of whether or not she participated. This setup will not always generate problematic consequences: provided the agent’s contribution makes a significant difference to the probability of task completion, the expected benefit of participating can still outweigh the expected cost. The trouble is that, in real-life cases where collective action seems vitally important, the contribution of any given individual to the probability of task success is often extremely small. The unpalatable result is that, as long as the cost attached to participation is

<sup>7</sup> Redundancy is also present to a significant degree within the genome, but I will not pursue this interesting analogy here (see, e.g., Thomas 1993, Nowak et al. 1997).

significant, agents have an incentive to avoid participating in these tasks, even when the benefit of task completion would be very large (see Olson 1965; Kagel and Roth 1995; Hardin 1971, 1982; Medina 2007).<sup>8</sup>

Thankfully, collective action still happens. The puzzle for game theory is to explain why. Voter turnout is a famous problem case. As Anthony Downs (1957) notes, the probability that one's vote will swing the outcome of an election is typically minuscule. Hence, even if the benefits of tipping the outcome would be very high, the *expected* benefit of voting will be outweighed by even a small cost. Downs emphasized the costs of following the election closely enough to make an informed choice, but even the utility costs involved in taking the time to vote may exceed the expected benefit. Yet voters do vote, often in large numbers. This mismatch between theory and data has proved an enduring thorn in the side for social scientists, and the steady stream of proposed solutions to the 'paradox of voting' continues to this day (see Feddersen 2004 and Dowding 2005 for reviews; see Guerrero 2010 for a recent proposal).

The redundancy we find in complex insect societies and multicellular organisms leads to a closely analogous puzzle. When redundancy is extreme, the contribution of any given individual to the probability of task success is extremely small. The result is that, as long as the cost attached to participating is significant, apathy will yield a higher expected payoff than participation even when the potential benefits of task completion are large.

Evidently, however, insects in complex societies and somatic cells in multicellular organisms do participate in large-scale collective tasks. Moreover, as we noted above, their efforts often amount to an extreme and non-reciprocal form of altruism: by developing as workers or somatic cells, they wholly forego any opportunity to transmit their genes to the next generation by a direct route. The challenge is to explain how this situation can arise in the course of an evolutionary transition. Why do agents sacrifice their direct fitness in order to contribute to large-scale collective action? How does natural selection solve the collective action problem?

The significance of this puzzle in the present context is that indirect fitness effects do not make it disappear. Adding indirect fitness effects to the payoff functions makes the benefit of task completion larger, from the point of view of the focal agent, than it would otherwise be. But when the difference a single contribution makes to the probability of task completion is tiny enough, the expected benefit of participating can be outweighed by even a small cost, even if the total benefit conferred by task completion is large.

#### A formal statement of the puzzle

It is easier to grasp the precise nature of the puzzle with the aid of a game-theoretic treatment.<sup>9</sup> Consider a simple case of task-based cooperation, in which the inclusive

<sup>8</sup> A collective action problem, thus construed, is equivalent to a public goods dilemma (or 'free-rider problem') in which the 'public good' is, somewhat counterintuitively, the probability that the relevant task is completed.

<sup>9</sup> In this subsection I draw on the formal treatment of collective action problems in Medina (2007).

fitness payoffs to the focal agent of participating ( $W_P$ ) and abstaining ( $W_A$ ) depend on the inclusive fitness benefit to that agent if the task is completed ( $B$ ), the cost of participating ( $c$ ), the group size ( $n$ ), the fraction ( $\gamma$ ) of group members *other than the focal agent* who participate, and the function ( $f$ ) relating the probability of task completion to the *total* fraction of group members who participate:<sup>10</sup>

$$W_P = B \cdot f(\gamma + 1/n) - c \quad (1)$$

$$W_A = B \cdot f(\gamma) \quad (2)$$

In short, the agent has to choose whether or not to pay a cost  $c$  in order to add  $1/n$  units to  $\gamma$ , and thereby make a contribution to the probability of task completion. Let us say that a given level of overall participation,  $\gamma_0 + 1/n$ , is *sustainable* if and only if  $W_P(\gamma_0) \geq W_A(\gamma_0)$ . From (1) and (2), we can see that this will be the case if and only if:

$$B \cdot (f(\gamma_0 + 1/n) - f(\gamma_0)) - c \geq 0$$

The puzzle is to explain how this condition can be satisfied given that (1)  $c \gg 0$ , and (2)  $\gamma_0$  corresponds to a level of participation that far exceeds the minimum required for the completion of the task, such that  $f(\gamma_0 + 1/n) - f(\gamma_0) \approx 0$ . How can extreme levels of redundancy be sustainable? More bluntly: why do workers in complex societies bother, when their individual contributions count for so little? The logic of the problem is parallel to that of a standard collective action problem. It is a problem that still arises even if indirect fitness effects are included in the  $B$  term, and even if  $B$  is large.

To briefly review: workers in eusocial colonies (and somatic cells in multicellular organisms) incur huge fitness costs, often to the extent of sacrificing their direct fitness entirely. This can be explained by inclusive fitness considerations on the assumption that the expected indirect benefit of cooperating exceeds the direct cost. The extreme redundancy complex societies contain seems to threaten this assumption, because it implies that the expected indirect benefit an individual agent receives in return for participating in a large-scale collective task is typically very small. Individual workers and somatic cells thus appear to incur huge direct costs in return for extremely small expected indirect benefits. This is a puzzle for the inclusive fitness paradigm.

### Solving the puzzle: the crane of coercion

There is no reason to suppose that the puzzle is insoluble. Indeed, I think we can solve it, and solve it within the constraints of inclusive fitness theory. I suggest, however, that the puzzle of extreme redundancy does indicate something interesting about the *order* in which the proximate aspects of social complexity arise in the

<sup>10</sup> In complex societies, the probability of task completion is likely to depend on numerous variables, not just on the overall degree of participation. For current purposes, however, we can assume that all these variables are held fixed, so that the overall degree of participation is the only factor that influences the probability of success. For simplicity, I treat  $B$  and  $c$  as constants, but in reality both may vary between agents.

course of a fraternal transition. For it gives us reason to suspect that extreme redundancy is unlikely to evolve as long as the personal fitness difference between participating in and abstaining from large-scale collective tasks remains significantly greater than zero. Recall the structure of the argument: when redundancy is extreme, a single worker (or somatic cell) makes no significant difference to the probability of task completion. Hence, as long as there is a significant fitness cost attached to participating, the worker will have an incentive to abstain, even if the potential benefit of task completion is large.

To solve the problem, we need a further mechanism that negates the personal fitness difference between participating and abstaining, and hence eliminates an agent's incentive to abstain even when its contribution to the probability of task completion is extremely small. We see a plausible mechanism in the eusocial Hymenoptera in the form of *worker policing*, whereby, in some colonies, around 98% of worker-laid eggs are eaten by the queen or by other workers (see Ratnieks and Wenseleers 2008). Both direct and indirect fitness considerations suggest plausible explanations for the evolution of policing. Policing provides a direct benefit, since the eggs nourish the workers who eat them. Moreover, it plausibly yields an indirect fitness benefit too, since haplodiploidy ensures that workers are more closely related to the offspring of the queen than to the offspring of their fellow workers—and so stand to gain in inclusive fitness terms from helping the queen to reproduce at other workers' expense (Ratnieks 1988).<sup>11</sup>

There are two ways to interpret the way in which policing transforms the collective action problem of task-based cooperation. We could say that the effect of policing is to reduce the cost of participation ( $c$ ), because workers have so little expected direct fitness anyway that they stand to lose relatively little by contributing to collective tasks. We could alternatively say that, though  $c$  remains high, policing imposes a selective cost ( $d$ ) on those who attempt to secure genetic representation in the next generation by a direct route. My own preference is for the second option: participating in collective tasks intuitively entails a cost to the viability and fecundity of a worker, and the effect of policing is to impose a counterbalancing cost on those who attempt to defect. On this interpretation, policing leads to the modified payoff functions:

$$W_P = B \cdot f(\gamma + 1/n) - c \quad (3)$$

$$W_A = B \cdot f(\gamma) - d \quad (4)$$

If  $c = d$ , and policing wholly counterbalances the cost of participating in collective action, then  $W_P > W_A$  even if  $f(\gamma + 1/n) - f(\gamma) \approx 0$ . High levels of participation will be sustainable, and there will be an incentive to participate even when redundancy is extreme. I suggest, moreover, that  $c = d$  is a reasonably plausible interpretation of the scenario we see in societies with highly effective coercive regimes. In such societies, coercion counterbalances the cost of

<sup>11</sup> When individuals are unrelated and policing behaviour imposes a cost on the actor, there is an incentive not to police: the result is a 'second-order free-rider problem' (Heckathorn 1989). Though this presents a major difficulty in human contexts, such a problem seems unlikely to arise in the context of the social insects, since policing plausibly confers both direct and indirect fitness benefits.

participation to such an extent that the chance a worker has of getting its genes into the next generation by a direct route is unaffected by whether or not it chooses to participate in collective tasks: either way, the chance is effectively zero. These, I suggest, are the ecological conditions in which extreme redundancy is likely to arise.

While the role of policing in resolving reproductive conflict in insect societies has been well documented in recent literature (see Ratnieks et al. 2006 for a review), this hypothesis highlights the potential *creative* role of coercion as an enabler for the evolution of important features of social complexity—aspects which, on the face of it, have little to do with the policing of egg-laying, but which would not be stable in its absence. When an effective coercive regime is in place, large numbers of workers may stably participate in large-scale cooperative tasks even when the expected inclusive fitness benefit conferred by their own individual contributions is extremely small, because the cost of participating is fully counterbalanced by the costs imposed on defectors. The effect of systematic coercion is thus to modify the selective environment in such a way as to make the evolution of extreme redundancy possible. Extreme redundancy, for its part, then helps enable the evolution of the extreme specialization characteristic of transitions in individuality (see ‘Why redundancy matters’).

If this hypothesis is on the right lines, the story that emerges is one in which indirect fitness benefits are indeed the driving force in the evolution of extreme redundancy and specialization, but in which they are not the *only* factor that matters. Instead, we have a four-step mechanism (Fig. 4). In the first step, kin selection favours coercive behaviour among the workers, imposing a cost on those who attempt to reproduce. Coercion results in a selective environment that is extremely hostile to any attempt by a worker to increase its inclusive fitness by a direct route; by contrast, a behaviour that increases inclusive fitness by an *indirect* route will be favoured, even when the expected benefit it confers is extremely small. This second step is, in a manner of speaking, a form of *niche construction* (sensu Odling-Smee et al. 2003), but it is the *social* environment rather than the natural environment that has been altered.<sup>12</sup> In the third step, further kin selection in the modified selective environment generates complex task-structures characterized by extreme redundancy. Finally, extreme redundancy enables the loss of latent versatility for the sake of increased specialization.

I have told this story in the specific context of complex sociality in the eusocial Hymenoptera. Could a similar mechanism explain the evolution of extreme redundancy and specialization in multicellular organisms? As Reeve and Jeanne (2003, p. 1044) note, mutations during mitotic cell division can generate considerable genetic heterogeneity in multicellular organisms, and this heterogeneity implies scope for reproductive conflict (see also Buss 1987; Michod 1999; Bourke 2011). Reeve and Jeanne proceed to argue that cell lines will have an incentive to coerce one another, leading to a scenario in which reproductive capacity is largely concentrated in a single dominant cell line. This will be the

<sup>12</sup> I borrow the notion of “social niche construction” from Powers et al. (2011) who apply it in a rather different context.

### Kin-selected coercion

*Workers are more closely related to the offspring of the queen than to the offspring of their co-workers (Ratnieks 1988); possible analogue in multicellular organisms (Reeve and Jeanne 2003).*

### Altered selective regime

*The selective environment is now extremely hostile to any attempt by workers to increase their fitness by a direct route.*

### Extreme specialization

*Extreme redundancy ensures robustness without the need for latent versatility. Latent versatility may be sacrificed for increased specialization.*

### Kin-selected redundancy

*There is now an incentive for workers to participate even when their individual contributions are insignificant.*

**Fig. 4** The “crane of coercion” hypothesis

slowest-dividing cell line, since this is the cell line to which the others have the highest mean relatedness. The upshot is an enforced germ-soma separation in which the direct fitness of the somatic cell lines is highly constrained; apoptosis among somatic cells may subsequently evolve as a form of self-restraint. We see a potentially analogous form of self-restraint in the eusocial insects, where workers, despite being able to activate their ovaries and produce male eggs, tend to do so *extremely* rarely: this restraint may plausibly have evolved as an adaptive response to the near impossibility of successfully protecting one’s eggs (Ratnieks and Wenseleers 2008). The extension of the hypothesis to origins of multicellularity is highly speculative, and there is evidently a need for further theoretical and empirical investigation. Nevertheless, we can conjecture that the presence of a coercive regime may be of general importance in driving fraternal transitions.

## The outlook for the inclusive fitness paradigm

Much of the foregoing discussion is highly sympathetic to the view that inclusive fitness theory has the resources to make sense of the major transitions. In ‘Proximate aspects of social complexity’, we canvassed several core features of complex cooperation. Two general morals leap out from this survey. First, many of the organizational features of the most complex insect societies are shared, in some form or another, by paradigm multicellular organisms. Second, the same features are also exhibited, albeit to a much lesser degree, by many simpler animal societies. The overall picture is one in which the same broad types of complex social phenomena recur throughout the biological hierarchy, wherever groups of entities are bound into stable, integrated wholes. Yet these features also generate new explanatory challenges for the study of social evolution.

In the second half the paper, we considered in detail one aspect of social complexity (namely, extreme redundancy) and the distinctive explanatory challenge it presents (namely, a biological analogue of the ‘paradox of voting’). The puzzle of extreme redundancy threatens an overly ambitious view according to which



transitions in individuality are held to present no special problems at all for inclusive fitness theory, but it seems doubtful that even the keenest advocates of the theory would seriously defend so strong a position. The “crane of coercion” hypothesis is consonant with a more modest outlook, on which kin selection is indeed pivotal in driving the evolution of extreme redundancy and the extreme specialization it may facilitate, but only as part of a multi-step process, and only in conjunction with a process of social niche construction.

The general moral, I suggest, should be one of qualified optimism. Inclusive fitness theory may well turn out to explain far more than it was originally designed to explain. If the case of extreme redundancy is indicative of a broader trend, however, it will do so only as part of an integrated theoretical approach that recognizes both the special problems posed by transitions in individuality and the distinctive mechanisms by which these problems may be overcome.

**Acknowledgments** I thank Tim Lewens, Kevin Brosnan, Andrew Bourke, Brett Calcott, Ellen Clarke, Patrick Forber, Peter Godfrey-Smith, Andrew Hamilton, Rufus Johnstone, Ben Kerr, Arnon Levy, Jason Noble, Samir Okasha, Cedric Paternotte, Paul Ryan, Denis Walsh, Joeri Witteveen and an anonymous referee for helpful comments and discussion, and I thank audiences at the IHPST (Paris), the University of Bristol and the University of Utah. This work was supported by the Arts and Humanities Research Council.

## References

- Anderson C, Franks NR (2001) Teams in animal societies. *Behav Ecol* 12:534–540
- Anderson C, McShea DW (2001) Individual versus social complexity, with particular reference to ant colonies. *Biol Rev* 76:211–237
- Anderson C, Franks NR, McShea DW (2001) The complexity and hierarchical structure of tasks in insect societies. *Anim Behav* 62:643–651
- Bonner JT (1988) *The evolution of complexity by means of natural selection*. Princeton University Press, Princeton
- Bourke AFG (2011) *Principles of social evolution*. Oxford University Press, Oxford
- Bourke AFG, Franks NR (1995) *Social evolution in ants*. Princeton University Press, Princeton
- Buss LW (1987) *The evolution of individuality*. Princeton University Press, Princeton
- Calcott B (2008) The other cooperation problem: generating benefit. *Biol Philos* 23:179–203
- Clarke E (2011) Plant individuality and multilevel selection theory. In: Calcott B, Sterelny K (eds) *The major transitions in evolution revisited*. MIT Press, Cambridge, pp 227–250
- Dowding K (2005) Is it rational to vote? Five types of answer and a suggestion. *Brit J Polit Int Rel* 7:442–459
- Downs A (1957) *An economic theory of democracy*. Harper and Row, New York
- Feddersen TJ (2004) Rational choice theory and the paradox of not voting. *J Econ Perspect* 18:99–112
- Frank SA (1994) Kin selection and virulence in the evolution of protocells and parasites. *Proc R Soc B* 258:153–161
- Gardner A, West SA, Wild G (2011) The genetical theory of kin selection. *J Evol Biol* 24(5):1020–1043
- Godfrey-Smith P (2009) *Darwinian populations and natural selection*. Oxford University Press, New York
- Gorelick R, Bertram SM, Killeen PR, Fewell JH (2004) Normalized mutual entropy in biology: quantifying division of labour. *Am Nat* 164:677–682
- Guerrero AA (2010) The paradox of voting and the ethics of political representation. *Philos Public Aff* 38:272–306
- Hamilton WD (1964) The genetical evolution of social behaviour. *J Theor Biol* 7:1–52
- Hamilton WD (1970) Selfish and spiteful behaviour in an evolutionary model. *Nature* 228:1218–1220
- Hardin R (1971) Collective action as an agreeable n-prisoners’ dilemma. *Behav Sci* 16:472–481
- Hardin R (1982) *Collective action*. Johns Hopkins University Press, Baltimore

- Heckathorn DD (1989) Collective action and the second-order free-rider problem. *Ration Soc* 1:78–100
- Herbers JM (1981) Reliability theory and foraging by ants. *J Theor Biol* 89:175–189
- Hölldobler B, Wilson EO (1990) The ants. Harvard University Press, Cambridge
- Kagel JH, Roth AE (1995) The handbook of experimental economics. Princeton University Press, Princeton
- Maynard Smith J (1982) Evolution and the theory of games. Cambridge University Press, Cambridge
- Maynard Smith J, Szathmáry E (1995) The major transitions in evolution. Oxford University Press, Oxford
- Medina LF (2007) A unified theory of collective action and social change. University of Michigan Press, Ann Arbor
- Michod RE (1983) Population biology of the first replicators: on the origin of the genotype, phenotype and organism. *Am Zool* 23:5–14
- Michod RE (1999) Darwinian dynamics: evolutionary transitions in fitness and individuality. Princeton University Press, Princeton
- Michod RE (2007) Evolution of individuality during the transition from unicellular to multicellular life. *Proc Natl Acad Sci USA* 104:8613–8618
- Michod RE, Herron MD (2006) Cooperation and conflict during evolutionary transitions in individuality. *J Evol Biol* 19:1406–1409
- Mirenda JT, Vinson SB (1981) Division of labour and specification of castes in the red imported fire ant *Solenopsis invicta* Buren. *Anim Behav* 29:411–420
- Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314:1560–1563
- Nowak MA, Boerlijst MC, Cooke J, Maynard Smith J (1997) Evolution of genetic redundancy. *Nature* 388:167–171
- Odling-Smee JF, Laland KN, Feldman MW (2003) Niche construction: the neglected process in evolution. Princeton University Press, Princeton
- Olson M (1965) The logic of collective action: public goods and the theory of groups. Harvard University Press, Cambridge
- Oster GF, Wilson EO (1978) Caste and ecology in the social insects. Princeton University Press, Princeton
- Powers ST, Penn AS, Watson RA (2011) The concurrent evolution of cooperation and the population structures that support it. *Evolution* 65:1527–1543
- Queller DC (1997) Cooperators since life began. *Q Rev Biol* 72:184–188
- Queller DC (2000) Relatedness and the fraternal major transitions. *Philos Trans R Soc B* 364:3142–3155
- Queller DC, Strassmann JE (1998) Kin selection and social insects. *Bioscience* 48:165–175
- Queller DC, Strassmann JE (2009) Beyond sociality: the evolution of organismality. *Philos Trans R Soc B* 355:1647–1655
- Ratnieks FLW (1988) Reproductive harmony via mutual policing by workers in eusocial Hymenoptera. *Am Nat* 132:217–236
- Ratnieks FLW, Anderson C (1999) Task partitioning in insect societies. *Insect Soc* 46:95–108
- Ratnieks FLW, Wenseleers T (2008) Altruism in insect societies and beyond: voluntary or enforced? *Trends Ecol Evol* 23:45–52
- Ratnieks FLW, Foster KR, Wenseleers T (2006) Conflict resolution in insect societies. *Annu Rev Entomol* 51:581–608
- Reeve HK, Jeanne RL (2003) From individual control to majority rule: extending transactional models of reproductive skew in animal societies. *Proc R Soc B* 270:1041–1045
- Strassmann JE, Queller DC (2007) Insect societies as divided organisms: the complexities of purpose and cross-purpose. *Proc Natl Acad Sci USA* 104:8619–8626
- Strassmann JE, Queller DC (2010) The social organism: congresses, parties and committees. *Evolution* 64:605–616
- Szathmáry E, Maynard Smith J (1997) From replicators to reproducers: the first major transitions leading to life. *J Theor Biol* 187:555–571
- Taylor C, Nowak MA (2007) Transforming the dilemma. *Evolution* 61:2281–2292
- Thomas JH (1993) Thinking about genetic redundancy. *Trends Genet* 9:395–399
- Trivers RL (1985) Social evolution. Benjamin/Cummings, Menlo Park
- West SA, Griffin AS, Gardner A (2007) Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J Evol Biol* 20:415–432
- West-Eberhard MJ (1975) The evolution of social behavior by kin selection. *Q Rev Biol* 50:1–33