Evolution of heritable behavioural differences in a model of social division of labour

Zsóka Vásárhelyi[1], Géza Meszéna[2] and István Scheuring[3]

[1] Department of Plant Systematics, Ecology and Theoretical Biology, Eötvös Loránd University, Budapest, Hungary
[2] Department of Biological Physics, Eötvös Loránd University, Budapest, Hungary
[3] MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, Eötvös Loránd University and the Hungarian Academy of Sciences, Budapest, Hungary

ABSTRACT

The spectacular diversity of personality and behaviour of animals and humans has evoked many hypotheses intended to explain its developmental and evolutionary background. Although the list of the possible contributing mechanisms seems long, we propose that an underemphasised explanation is the division of labour creating negative frequency dependent selection. We use analytical and numerical models of social division of labour to show how selection can create consistent and heritable behavioural differences in a population, where randomly sampled individuals solve a collective task together. We assume that the collective task needs collaboration of individuals performing one of the two possible subtasks. The total benefit of the group is highest when the ratio of different subtasks is closest to 1. The probability of choosing one of the two costly subtasks and the costs assigned to them are under selection. By using adaptive dynamics we show that if a trade-off between the costs of the subtasks is strong enough, then evolution leads to coexistence of specialized individuals performing one of the subtasks with high probability and low cost. Our analytical results were verified and extended by numerical simulations.

INTRODUCTION

Behavioural differences among conspecifics of the same population have been reported from all kinds of living organisms, from bacteria to vertebrates, including humans (*Dall, Houston & McNamara, 2004*; *Dingemanse et al., 2004*; *Stampsa, Briffa & Biro, 2012*; *Grinsted et al., 2013*; *Cordero & Polz, 2014*). A bacterium's behaviour is determined by its possession or lack of certain genes. However, as organismal complexity rises, behaviour also becomes more subtle and less predictable, reaching the greatest complexity in humans. Interestingly, while variation in bacterial behaviour is widely studied as a classical model of fluctuating or negative frequency dependent selection (*Cordero & Polz, 2014*), the adaptive value of individual variation in animal personality, especially in humans, is still the subject of debate (*Penke, Denissen & Miller, 2007*; *Montiglio, Ferrari & Réale, 2013*).

There are several approaches to explaining the ultimate causes of human and animal personality variation (where animal personality is also called as coping style, behavioural syndrome, etc. (*Wolf et al., 2007*)), and though the available methods and background knowledge are quite different in the two areas, the basic explanations are rather similar. These include trade-offs between different life-history strategies (i.e., reproducing early vs. growing more) (*Wolf et al., 2007*) or personality dimensions (*Nettle, 2006*), individual niche specialisation (*Bergmüller & Taborsky, 2010*; *Montiglio, Ferrari & Réale, 2013*) and different forms of fluctuating or frequency dependent selection (*Penke, Denissen & Miller, 2007*).

However, the possibility that selection for cooperation can also play a role in causing or maintaining this diversity has not yet received proper attention. We think that there are many social situations with dividable tasks that require mutualistic cooperation; that is, cooperation from which all participants benefit, where behaviourally diverse groups must be better off. Examples for such dividable tasks could be the operation of hives, webs, etc. (*Buech, 1995*; *Duarte et al., 2011*; *Grinsted et al., 2013*), the rearing of off-spring (*Clutton-Brock, Russell & Sharpe, 2004*; *Barta et al., 2014*), group hunting (*Stander, 1992*; *Newton-Fisher, 2014*) or even trade in human context (*Gowdy & Krall, 2013*).

We propose that division of labour between socially related conspecifics can cause and maintain variation in behaviour through negative frequency-dependent selection that continuously decreases a phenotype's benefit as its frequency increases. Selection thus creates diversity, by favouring the rare, acting on preferences and skills; that is, on personality traits. The aim of this article is to study such a strategic model of division of labour. But first, because the theoretical and empirical literature on division of labour (henceforth DL) is so diverse, we find it useful to give a short outline and classification of the phenomena.

DL, which is a cooperative act where the acting parties are engaged in different tasks, can be observed on every level of organization from genes to social groups (*Maynard Smith & Szathmáry, 1995*). However, the literature is somewhat confusing in what DL actually is, as authors put the same label on very distinct phenomena or mechanisms. Therefore, we will categorize DL along two dimensions: whether it is *linked to reproduction* or is *unlinked to reproduction*, and whether it occurs *within* an individual or *between* individuals (Table 1.). *Reproduction-linked* DL means that it directly arises from some kind of reproductive DL. Sensu stricto reproductive DL means that some members of a group are reproducing, while others are not (*Simpson, 2011*) like the sterile castes in eusocial insect societies or the somatic cells after the germ-soma differentiation in multicellular organisms. We think that reproductive DL in a broader sense happens between other-sex individuals too. These could allow reproduction-linked DL to arise, such as the functional and physical specialisation of reproductively inactive female ants and Trematoda larvae to defence (*Duarte et al., 2011*; *Lloyd & Poulin, 2014*), and also any sexually breeding pair of parents taking mildly or strongly different roles in rearing their offspring (*Buech, 1995*; *Barta et al., 2014*; *Gurven et al., 2009*).

Many examples of DL also occur between reproductively similar (or nearly identical) members of a group and therefore are not the direct result of reproductive asymmetries:

Table 1 A possible classification of the diverse phenomena labelled as division of labour (DL). DL can occur within or between individuals and can be linked to reproduction or unlinked, that is independent of it (for further details see the main text). This paper concentrates on the section with grey background that we call social DL.

	Reproduction-linked	Reproduction-unlinked
Within ind.	– Germ-soma differentiation in Volvox colonies (*Michod et al., 2006*) – Non-reproductive polyps in Hydrozoa colonies (*Cartwright, 2003*)	– Paralogous genes (*Rueffler, Hermisson & Wagner, 2012*) – Enzymes in a protocell (*Boza et al., 2014*) – Specialized limbs (*Rueffler, Hermisson & Wagner, 2012*)
Between ind.	– Sterile workers in Hymenoptera societies (*Sherman et al., 1995*) – Sterile Trematoda rediae (*Lloyd & Poulin, 2014*) – Task specialisation due to the reproductive hierarchy in meerkats (*Clutton-Brock, Russell & Sharpe, 2004*) – Biparental care in different animal societies (*Barta et al., 2014*) – Sexual DL in human societies (*Gurven et al., 2009*)	– Cooperative hunting in chimpanzees (*Boesch, 2002*) – DL between reproductive female spiders in a common web (*Grinsted et al., 2013*; *Wright, Holbrook & Pruitt, 2014*) – Task specialization of sterile Hymenoptera workers (*Duarte et al., 2011*)

we call this *reproduction-unlinked DL*. Examples include complementary enzymes in a protocell (*Boza et al., 2014*), the specialisation of limbs or organs (*Rueffler, Hermisson & Wagner, 2012*), task-specialisation in social spider colonies (*Wright, Holbrook & Pruitt, 2014*) and Hydrozoa colonies (*Cartwright, 2003*), and the cooperative hunting of lions and chimpanzees (*Stander, 1992*; *Newton-Fisher, 2014*).

As the abovementioned examples show, DL is not restricted to groups of individuals: any cooperatively linked pair or group of any kind of organisation (i.e., organelles, cells, tissues, etc.) can be involved. Hence, we differentiate between *within* and *between* individuals (Table 1). Naturally, boundaries between the categories outlined above and in Table 1 are sometimes blurred and controversial, yet we think that such a division still helps to see differences and similarities between these interconnected phenomena.

Each of the four categories circumscribes a type of DL that has been studied on its own (for examples and references see Table 1.), but one can also unfold a bigger picture with several interrelations between categories. For instance, the process by which *between individuals* DL becomes *within individual* DL is where a higher level or organisation emerges from a cooperative group of lower level organisms (or organizations). This mechanism often contributes to major transitions in evolution (*Maynard Smith & Szathmáry, 1995*). Classic examples are organelles constituting the eukaryota cell, the differentiation of germ and soma in multicellular animals and the appearance of sterile castes in eusocial insect societies, which are known as superorganisms (*Maynard Smith & Szathmáry, 1995*). Eusocial insect societies do not just represent a major evolutionary transition, but are also good examples for the blurred boundaries between the divisions outlined above, as they can be seen as one superorganism with an incomplete germ-soma differentiation and a flexible functional specialisation or as a social group of insects with strict reproductive DL and highly evolved DL among the non-reproductive members of the group with both physical and functional specialisation. Therefore, they can be put in both rows, that is in all elements of Table 1.

In this paper we are interested in *reproduction-unlinked* DL *between* individuals, which we will call *social* DL. We want to make a general argument about the possible role of social DL in creating and maintaining personality variation, as we find the literature lacking in this area. We present two variants of an analytically tractable model of DL. We show that skill differentiation is typical in both cases, even though they are not qualitatively nor quantitatively identical. Then we study an individual-based model representation of the problem to support and widen the analytical results, and finally we show how this mechanism can lead to diverse personalities or behavioural syndromes.

THE ANALYTICAL MODEL

Model definition

Our model studies a large well-mixed population, where members of small groups solve a collective task. Solving this task produces a common good for the group which is distributed among the members after performing the task. Groups consist of a fixed number of (N) individuals sampled randomly from the population. Individuals can choose between two kinds of subtasks, A and B, which we will also call strategies. When entering a group, individuals make their subtask choice before they observe the choices of others, and task switching is not allowed. Individuals can specialize more or less to one strategy, but then they necessarily loose some of their flexibility: the more an individual specializes on one subtask, the less is his cost for performing it and more his cost for performing the other subtask. Hence, the abilities to perform the two strategies are not independent, and are in negative trade-off.

Individuals are characterised by two heritable traits: the probability of choosing A or B strategies, and the degree of specialisation to A or B. For simplicity, we assume that the population is asexual, so progenies own the same trait values as their parent did. We apply the methodology of adaptive dynamics to analyse the course of evolution. Our main interest is whether DL emerges as a result of an evolutionary branching into two subpopulations specialized for the two subtasks.

Costs

Each individual is characterised by the propensities of choosing each subtask and the costs assigned to performing them. Costs c_A and c_B, associated with tasks A and B, respectively, are constrained by the trade-off

$$c_A^\alpha + c_B^\alpha = k^\alpha \tag{1}$$

(Fig. 1), where $k > 0$ represents an additional cost to a basic minimal cost (which is scaled to be zero in the model). $\alpha > 1$ means that the average cost is higher when individuals play mixed strategies (i.e., choose both subtasks with nonzero probability) than when they play only one (pure) strategy, A or B, with probability 1. The opposite is true when $\alpha < 1$. We will choose c_A as the independent variable, therefore

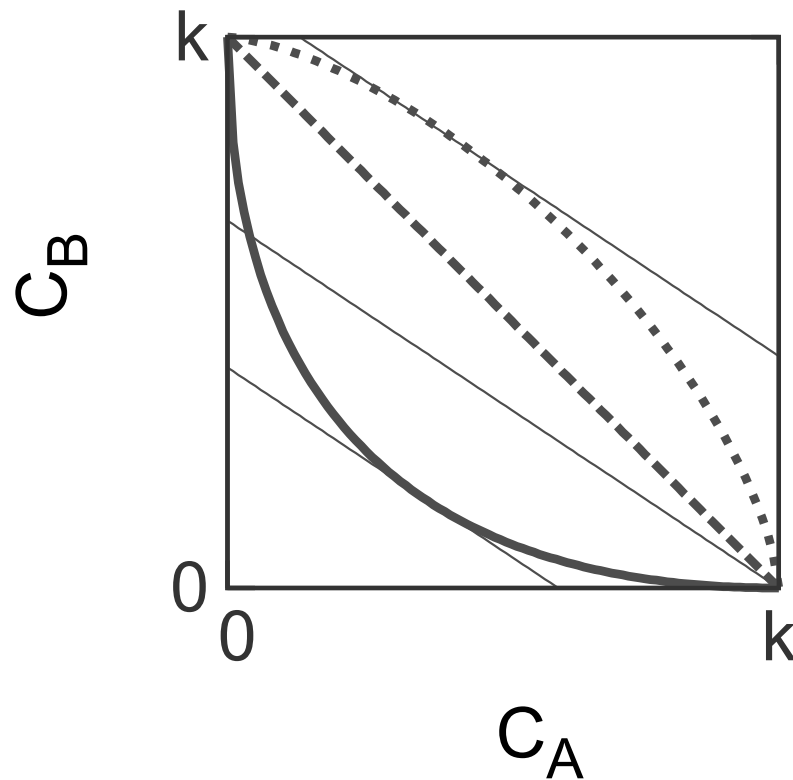$$c_B = \sqrt[\alpha]{k^\alpha - c_A^\alpha}. \tag{2}$$

Figure 1 The characteristic trade-offs and cost-minimisation. The thick lines are the trade-off curves for $\alpha = 0.5$ (solid), $\alpha = 1.0$ (dashed) and $\alpha = 1.5$ (dotted). In general, trade-off curves are convex for $\alpha < 1$ and concave for $\alpha > 1$. Thin lines represent iso-$\bar{c}$ lines of given q values (here $q = 0.4$). Higher the position of a line, the larger average cost, $\bar{c}$, it means. As we are looking for the lowest $\bar{c}$, we look for the lowest points where these iso-$\bar{c}$ lines touch the trade-off curves at least in one point. For $\alpha < 1$, we get the minimal $\bar{c}$ where the iso-$\bar{c}$ line is a tangent of the concave trade-off curve, the coordinates of the one common point define the optimal c_A and c_B. The maximal extremum we get where the iso-$\bar{c}$ line is a tangent of the convex trade-off curve. For $\alpha \geq 1$, the minimum is one of the endpoints of the scale.

Observe that the derivative of this dependence at the two extreme points is either zero or infinite for $\alpha \neq 1$.

Benefits

The total benefit of a group, $\beta(i)$, depends on the number of members (i) playing A (and consequently the number of members ($N - i$) playing B). $\beta(i)$ is a function typically with one maximum in $\{0, 1, ..N\}$. That is, there exists a distribution of A and B tasks where solving the collective task brings the highest possible benefit. We assume that β is symmetrical, i.e.,

$$\beta(N - i) = \beta(i), \tag{3}$$

and has a single maximum in the middle.

We will study two types of models differing only in their way of benefit sharing. In model (I), the group benefit, $\beta(i)$, is shared out equally among the group members, which could be a model of a cooperatively acquired common good from which nobody can be

Vásárhelyi et al. (2015), *PeerJ*, DOI 10.7717/peerj.977

5/22

excluded. In model (II), $\beta(i)$ is first split into two and then, within the subgroups playing A or B, halves of it are shared out equally, representing a sharing system, where the scarcer strategy gives a higher return. Therefore the individual benefits in the subgroups A and B are

$$b_A(i) = \begin{cases} \beta(i)/N & \text{for model (I)} \\ \frac{1}{2}\beta(i)/i & \text{for model (II)} \end{cases} \tag{4}$$

$$b_B(i) = \begin{cases} \beta(i)/N & \text{for model (I)} \\ \frac{1}{2}\beta(i)/(N-i) & \text{for model (II).} \end{cases} \tag{5}$$

We discuss these two models together until differences emerge. In the following, we will dismiss the term 1/2 from the model (II) calculations, as it makes only a quantitative difference. Note that

$$\Delta b(i) = b_A(i+1) - b_B(i) \tag{6}$$

is the benefit advantage (or disadvantage) of playing strategy A over strategy B for an individual, provided that the rest of the group have i members playing A. It is easy to see that

$$\Delta b(i) = -\Delta b(N-1-i), \tag{7}$$

which we will use later on.

Payoffs

The two *independent* heritable traits of individuals are $q \in [0,1]$, the probability of choosing subtask A, and $c_A \in [0,k]$, the cost assigned to this task. The expected payoff values for an individual playing both strategies are:

$$w_A(\bar{q}, c_A) = \sum_{i=0}^{N-1} f_{i,N-1}(\bar{q}) b_A(i+1) - c_A$$

$$w_B(\bar{q}, c_A) = \sum_{i=0}^{N-1} f_{i,N-1}(\bar{q}) b_B(i) - (k^\alpha - c_A^\alpha)^{\frac{1}{\alpha}}, \tag{8}$$

where $\bar{q}$ denotes the average q value of all the other population members, i.e., the probability of a randomly chosen individual playing A. The value for c_B was substituted. The probability of having i individuals playing A among the other $N-1$ group members is given by the binomial distribution

$$f_{i,N-1}(\bar{q}) = \binom{N-1}{i} \bar{q}^i (1-\bar{q})^{N-1-i} \tag{9}$$

obeying the symmetry relation

$$f_{i,N-1}(\bar{q}) = f_{N-i,N-1}(1-\bar{q}). \tag{10}$$

Vásárhelyi et al. (2015), *PeerJ*, DOI 10.7717/peerj.977

6/22

Finally, the expected payoff for an individual with trait values $\{q, c_A\}$ is

$$\overline{w}(q, c_A, \overline{q}) = q w_A(\overline{q}, c_A) + (1 - q) w_B(\overline{q}, c_A). \tag{11}$$

Cost optimisation

The sub-problem of cost-minimisation of an individual with a fixed q value can be considered separately, because it depends only on the q value of the individual and contributes to the payoff additively. Therefore, our first interest is to find the optimal c_A cost strategy that minimizes the average cost

$$\overline{c}(c_A) = q c_A + (1 - q) c_B = q c_A + (1 - q)(k^\alpha - c_A^\alpha)^{\frac{1}{\alpha}} \tag{12}$$

for a fixed q.

For $\alpha = 1$ it reduces to

$$\overline{c}(c_A) = q c_A + (1 - q)(k - c_A) = k(1 - q) + (2q - 1)c_A, \tag{13}$$

which is monotonous in c_A. For $q < \frac{1}{2}$ it means that cost minimization results in $c_A = k$ (full specialisation for task B), while $q > \frac{1}{2}$ makes $c_A = 0$ (full specialisation for task A) the optimal strategy. For $q = \frac{1}{2}$, c_A is a neutral trait, which does not affect the average cost.

For $\alpha \neq 1$ we can look for an internal extremum determined by

$$\frac{d\overline{c}(c_A)}{dc_A} = q - (1 - q)c_A^{\alpha-1}(k^\alpha - c_A^\alpha)^{(1-\alpha)/\alpha} = 0, \tag{14}$$

leading to

$$c_A^{\text{extr}}(q) = k\left[\left(\frac{q}{1-q}\right)^{\frac{\alpha}{1-\alpha}} + 1\right]^{-1/\alpha} \tag{15}$$

extreme c_A value. This function decreases for $\alpha < 1$ and increases for $\alpha > 1$. Relatedly, as it can be seen in Fig. 1, this extremum is the cost minimum for $\alpha < 1$ and the cost maximum for $\alpha > 1$.

Hence, for $\alpha < 1$, the cost optimum is internal and is determined by Eq. (15), while for $\alpha \geq 1$, it is at one of the end-points according to

$$c_A^{\min}(q) = \begin{cases} 0, & \text{for } q > \dfrac{1}{2} \\ k, & \text{for } q < \dfrac{1}{2}. \end{cases} \tag{16}$$

In the latter case the two end points are equally optimal for $q = \frac{1}{2}$.

The cost optimisation problem is trivial: $q = 0$, or 1, when the individual always choose the same subtask. Then it is optimal to fully specialize to the same subtask, leading to zero cost, and pessimal to specialize to the other subtask, leading to the maximal cost k. If q changes from 0 to 1, the optimal c_A changes from k to 0. This transition is continuous for

$\alpha < 1$, but abrupt at $q = \frac{1}{2}$ for $\alpha > 1$. It can be seen that the optimal $\bar{c}$ increases with q until $q = \frac{1}{2}$ and decreases afterwards.

Singular points, invasion analysis

Now we turn our attention to the full evolutionary problem, the invasion of a mutant strategy and the emergence of DL, which is inherently frequency-dependent, as the benefit of an individual no longer depends only on the traits of its own, but on the traits of its opponents too. Therefore, we apply the methodology of adaptive dynamics (*Geritz et al., 1997*). In particular, we consider pairwise invasion, because it is known that continuous evolution (i.e., evolution via small muational steps) is controlled by invasion (*Meszéna et al., 2005*; *Geritz et al., 2002*; *Geritz, 2005*).

We describe an individual with the vector x: $x = \{q, c_A\}$ and assume that the population has two types of members, a common (resident) and a rare (mutant) type. Denoting the mutant traits by prime marks, the relative fitness of the mutant is

$$W(x', x) = \overline{w}(q', c_A', q) - \overline{w}(q, c_A, q), \tag{17}$$

which is 0 by definition for $x' = x$. (Since the mutant is rare, it does not modify q in the leading term.)

Note that the fitness is linear in q', but is not so in q. Linearity in a rare trait is a common feature of game theoretic models with mixed strategies (*Hofbauer & Sigmund, 1998*), leading to non-standard behaviour relative to standard adaptive dynamics (*Meszéna et al., 2001*). On the other hand, the non-linearity in q is different from those commonly experienced in matrix games, and it is the consequence of the focal individual not playing against another individual, but instead, against a group of individuals.

Following the standard analysis we are interested in the singular points of the dynamics and their stabilities (*Geritz et al., 1997*; *Leimar, 2009*). x^* is a singular point, if

$$\left. \frac{\partial W(x', x)}{\partial x_i'} \right|_{x'=x} = 0, \tag{18}$$

i.e., if

$$\left. \frac{\partial W(x', x)}{\partial q'} \right|_{x'=x} = w_A(q, c_A) - w_B(q, c_A) = 0, \tag{19}$$

and if

$$\left. \frac{\partial W(x', x)}{\partial c_A'} \right|_{x'=x} = -\frac{d\bar{c}(c_A)}{dc_A} = 0. \tag{20}$$

We have already solved Eq. (20) by Eq. (15) for $\alpha \neq 1$. Substituting it into (19), q is determined by

$$G(q) + H(q) = 0 \tag{21}$$

at the singular point, where

$$G(q) = \sum_{i=0}^{N-1} f_{i,N-1}(q) \Delta b(i) \tag{22}$$

is the expected benefit advantage for the focal individual to play A, instead of B, and

$$H(q) = -\frac{d\bar{c}(c_A^{\text{extr}}(q))}{dq} = k\left[\left(\frac{q}{1-q}\right)^{\frac{1}{1-\alpha}} - 1\right]\left[\left(\frac{q}{1-q}\right)^{\frac{\alpha}{1-\alpha}} + 1\right]^{\frac{-1}{\alpha}} \tag{23}$$

is the contribution from the cost term, valid for $\alpha \neq 1$.

It follows from the symmetry of $\Delta b(i)$ and $\Delta f_{i,N-1}(q)$ (see (7) and (10)) that $G(q) = -G(1-q)$, so $G(1/2) = 0$ for model (I) and model (II) as well. One can see that $H(q)$ is also zero for $q = 1/2$; that is (19) is satisfied for $q = 1/2$ at $\alpha \neq 1$.

The case $\alpha = 1$ requires separate consideration, as the internal extremum of the average cost c_A does not exist in this case. However, recall that cost is a neutral trait for $\alpha = 1, q = \frac{1}{2}$. That is, the singular point is replaced by a singular line at $q = \frac{1}{2}$. This situation is usually neglected in adaptive dynamics.

Now we study the possibility of other singular points. It is easy to see that $G(q) > 0$ if $q < 1/2$, and consequently $G(q) < 0$ if $q > 1/2$. The same is true for function $H(q)$ for $\alpha > 1$, because $\bar{c}$ has a minimum at $q = 0$. Therefore, the only solution of $G(q) = 0$ is $q = 1/2$. Further, similar to $G(q)$, $H(q) = -H(1-q)$, and it can be shown that if $\alpha < 1$ ($\alpha \geq 1$), then $H(q)$ is strictly monotonously increasing (decreasing) with

$$\lim_{q \to 0} H(q) = -\theta k \tag{24}$$

$$\lim_{q \to 1} H(q) = \theta k, \tag{25}$$

where

$$\theta = \begin{cases} -1, & \text{for } \alpha > 1 \\ 0, & \text{for } \alpha = 1 \\ 1, & \text{for } \alpha < 1, \end{cases} \tag{26}$$

that is, for $\alpha = 1$ $H(q) \equiv 0$. Consequently, $q^* = 1/2$ is the only solution of (21) if $\alpha \geq 1$, so there is no other singular point except x^*. In contrast, for $\alpha < 1$ other singular points are possible too. It is easy to see that because of the continuity and the symmetry of the functions G and H, if

$$G(1) + H(1) > 0 \tag{27}$$

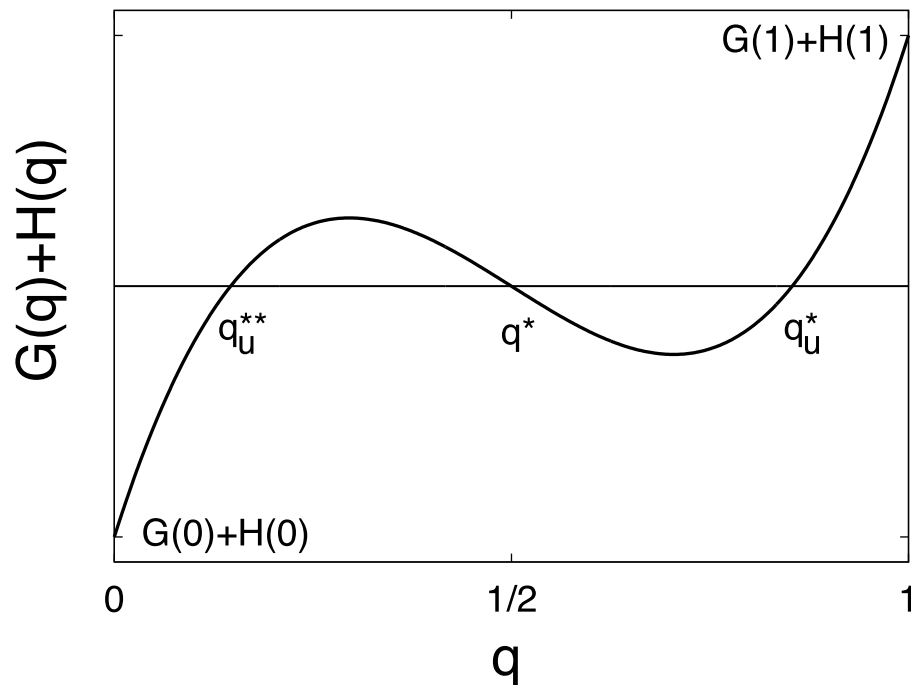$$\frac{d\left[G(q) + H(q)\right]}{dq}\bigg|_{q=1/2} < 0, \tag{28}$$

Figure 2 Function $G(q) + H(q)$ on a schematic figure with singular points x_u^{**}, x_u and x_u^*, when (Eq. (28)) are valid.

then at least one additional pair of singular points exists besides x^* (Fig. 2). Denote these points by $x_u^* = \{q_u^*, c_u^*\}$, where $q_u^* > 1/2$ and $x_u^{**} = \{q_u^{**}, c_u^{**}\}$, where $q_u^{**} = 1 - q_u^* < 1/2$, because of the symmetry of G and H. (If more than one such a point pair exists, then it is assumed that x^* has the biggest and x^{**} has the smallest equilibrium frequency.) Based on these results and assumptions, in the following subsection we investigate the invasion of a mixed strategy solving both subtasks in a resident population, where only one of the pure strategies is present as the resident strategy.

Invasion of a pure strategy by a mixed one

We consider the case when only one pure cost-free strategy is present initially; that is, the evolution begins either from the $x^{(1)} = \{1, 0\}$ or from the $x^{(0)} = \{0, k\}$ state, and we are interested in whether a mutant mixed strategy solving the alternative subtask with nonzero probability can invade or not. (Here, cost-free means that the subtask has no *additional* cost, since we can assume that any subtask has a constant basic cost which we can build in the benefit.) Without losing generality, let us look at the case of $x^{(1)}$, that is, every individual chooses subtask A and this subtask is cost-free. Now we study the invasion of a mutant in the neighborhood of $x^{(1)}$ by asking whether there are $x' = \{q', c_A'\}$ mutants with higher fitness than the resident one, that is whether

$$W(x', x^{(1)}) \approx \left.\frac{\partial W(x', x)}{\partial q'}\right|_{x'=x^{(1)}} \delta q + \left.\frac{\partial W(x', x)}{\partial c_A'}\right|_{x'=x^{(1)}} \delta c_A \tag{29}$$

can be positive, or not, where $\delta q = q' - q^{(1)} = q' - 1 \leq 0$ and $\delta c_A = c'_A - c_A^{(1)} = c'_A \geq 0$ in such a way that at least one of these terms differs from zero. It is clear that the partial derivatives become

$$\left.\frac{\partial W(x',x)}{\partial q'}\right|_{x'=x^{(1)}} = G(1) + k, \tag{30}$$

$$\left.\frac{\partial W(x',x)}{\partial c'_A}\right|_{x'=x^{(1)}} = -1. \tag{31}$$

The second term in (29) is always negative ($-1 \times \delta c_A < 0$). Thus $W(x', x^{(1)})$ can be postitive only if $G(1) + k$ is negative (since $\delta q < 0$), which is the condition for the invasion of a mutant (if $\delta c_A = 0$ then $q' < 1$ spreads). The sign of $G(1) + k$ depends on the parameters and the type of the model. Depending on the model type we have two cases:

$$G(1) + k = \begin{cases} [\beta(N) - \beta(N-1)]/N + k, & \text{for model (I)}, \\ \beta(N)/N - \beta(N-1) + k, & \text{for model (II)}. \end{cases} \tag{32}$$

Considering model (I) it is clear that $\beta(N) - \beta(N-1) < 0$, thus (32) can be negative. However, if N is large enough and/or the total cost k is high enough, then $G(1) + k > 0$ so $W(x', x^{(1)}) < 0$ and thus $x^{(1)}$ is resistant against the invasion of mixed strategies. The situation is a bit different for model (II). Again, $\beta(N)/N - \beta(N-1)$ is negative, but this value is smaller than in the previous case, thus $G(1) + k$ can easily remain negative if N is large and $\beta(N-1) > k$.

Because of the symmetry in the model, the condition for the invasion at $x^{(0)}$ is qualitatively the same as at the point $x^{(1)}$. Here if $G(0) - k > 0$, then a mixed strategy mutant can spread, which is more probable in model (II), as before. However, we must admit a mathematical nuance at this point. The trade-off function (2) has a zero-or-infinite derivative at $c_A = 0$, or 1, for $\alpha \neq 1$. That is, if mutation steps are uniformly small, when parametrised by c_A, they become non-uniform when re-parametrised by c_B. From a biological point of view, the natural assumption is that mutations are uniformly small in c_A in the vicinity of $c_A = 0$ and uniformly small in c_B in the vicinity of $c_B = 0$. Thus c_B, the cost of the mutant B strategy must be the basic variable at $x^{(0)}$.

Further progress after the initial invasion in the direction of decreasing q is very different for $\alpha < 1$ and for $\alpha > 1$. So let us assume that invasion is possible. Then a new mutant x' spreads and fixates in the population. Clearly $q' < 1$ and $c'_A \geq 0$ in this new resident state. To study the invasion of the next mutant x'' we can follow the same analysis as before. However, the partial derivative of the fitness difference as a function of c''_A is

$$\left.\frac{\partial W(x'',x')}{\partial c''_A}\right|_{x''=x'} = \begin{cases} -q' + (1-q')\mathcal{O}((c'_A)^{\alpha-1}), & \text{for } \alpha > 1, \\ 1 - 2q', & \text{for } \alpha = 1, \\ -q' + (1-q')/\mathcal{O}((c'_A)^{\alpha-1}), & \text{for } \alpha < 1 \end{cases} \tag{33}$$

(cf. (14)). If $\alpha < 1$ then this derivative is a large positive quantity for very small c_A. Therefore $c''_A > 0$ easily spreads, and the invasion process continues to the inner singular

point. Recall that the cost optimum for a given q is internal for $\alpha < 1$. Consequently, the "best" direction for a small mutation follows the direction of the curve $c_A^{\text{extr}}(q)$. That is, we have an essentially one-dimensional dynamics along this curve; direction is determined by the sign of $-q' \approx -1$, thus invasion of $c_A'' > 0$ is not supported and invasion events stop in x^m where $q^m < q^{m-1}$, $c_A^m = c_A^{m-1} = 0$. Because the cost optimum is always at one of the end points of the c_A scale, if the pure strategy is invaded, then evolution proceeds along the edge $c_A = 0$ of the strategy space. Direction of this one-dimensional strategy is determined by the sign of $G(q) + k$, as the cost strategy does not change; a rest point is reached when this variable becomes zero. As $G(\frac{1}{2}) = 0$, the rest point x^m is reached at $q > \frac{1}{2}$, i.e., before the cost optimum switches to the other extreme.

Convergence and branching of the singular point *x**

The question is the $x^* = \{q^*, c_A^*\}$ singular point's convergence stability and the conditions of branching. The convergence of a singular point means that in its neighborhood, evolution leads the system towards this singular point (*Geritz et al., 1997*). However, when close to it, branching gives rise to two different phenotypes in the evolutionary optimum. This is the state where coexistence of different types or DL is observed. Following *Leimar (2005)* and *Leimar (2009)*, to study convergence we have to find out whether the Jacobian matrix $J = H + Q$ of the selection gradient is positive or negative definite, where

$$H_{i,j} = \left.\frac{\partial^2 W(x',x)}{\partial x_i' \partial x_j'}\right|_{x'=x} \tag{34}$$

$$Q_{i,j} = \left.\frac{\partial^2 W(x',x)}{\partial x_i' \partial x_j}\right|_{x'=x}, \tag{35}$$

and x_i denotes the coordinates of the x vector. If $\mathbf{J}$ is negative definite, then x is strongly convergence stable. That is, this state is convergence stable independently of the characteristics of covariance among the mutations in x_1 and x_2. Further, if $\mathbf{H}$ is positive definite or indefinite, then the singular point is a branching point (*Leimar, 2009*).

It is easy to see that $H_{1,1} = Q_{1,2} = Q_{2,1} = Q_{2,2} = 0$ in our model, consequently the $\mathbf{J}$ matrix becomes

$$\begin{pmatrix} Q_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{pmatrix}. \tag{36}$$

We have seen that if $\beta(i)$ is symmetrical then $\{q^*, c_A^*\} = \{\frac{1}{2}, \frac{k}{\sqrt[\alpha]{2}}\}$ becomes a singular point. Then the $\mathbf{J}$ matrix becomes:

$$\begin{pmatrix} \Delta w(q^*) & -2 \\ -2 & \dfrac{\alpha - 1}{c_A^*} \end{pmatrix}, \tag{37}$$

where $\Delta w(q^*)$ is

$$\Delta w(q^*) = \frac{\partial [w_A(q, c_A) - w_B(q, c_A)]}{\partial q}\bigg|_{q=q^*, c_A=c_A^*}. \tag{38}$$

After the derivation we get

$$\Delta w(q^*) = \sum_{i=0}^{N-1} \binom{N-1}{i} q^{*i-1}(1-q^*)^{N-2-i}[i - (N-1)q^*]\Delta b(i). \tag{39}$$

Using $q^* = 1/2$, the symmetry of $\beta(i)$ and Eq. (22) we get

$$\Delta w(1/2) = \frac{1}{2^{N-3}} \sum_{i=0}^{N-1} \binom{N-1}{i} i\Delta b(i). \tag{40}$$

Since $\beta(i) = \beta(N-i)$, it is easy to see that $\Delta w(1/2) < 0$ for both models.

The $\{q^*, c_A^*\} = \{\frac{1}{2}, \frac{k}{\sqrt[\alpha]{2}}\}$ point is convergence stable (eigenvalues of (37) are negative) if

$$\Delta w(1/2) + \frac{\alpha - 1}{c_A^*} < 0$$

$$\Delta w(1/2)\frac{\alpha - 1}{c_A^*} - 4 > 0. \tag{41}$$

Since $\Delta w(1/2) < 0$, the second criterion can be true only if $\alpha < 1$. However, in this case, the first criterion will always be true. So the conditions for the singular point's convergence stability will be

$$\Delta w(1/2)\frac{2^{1/\alpha}(\alpha - 1)}{4} > k. \tag{42}$$

We note here that $\Delta w(1/2)$ is greater for model (I) than for model (II) at the same $\beta(i)$, thus the above criterion can be satisfied with a smaller α (stronger trade-off) in model (I) than in model (II). Similarly, $\Delta w(1/2)$ is greater if the maximal value of $\beta(i)$ is smaller, that is, branching needs stronger trade-off if the benefit to cost ratio of the collective task is smaller.

We can also get the **H** matrix using the formerly shown calculations:

$$\begin{pmatrix} 0 & -2 \\ -2 & \dfrac{\alpha - 1}{c_A^*} \end{pmatrix}, \tag{43}$$

It is easy to show that **H** is indefinite, therefore the singular point will always be a branching point. Obviously, the indefiniteness is related to the fact that the cost is minimal at the corners $\{1, 0\}$ and $\{0, k\}$ but maximal at $\{1, k\}$ and $\{0, 0\}$.

To summarize: the $\{q^*, c_A^*\} = \{\frac{1}{2}, \frac{k}{\sqrt[\alpha]{2}}\}$ singular point can only be convergence stable if $\alpha < 1$, which is fulfilled if condition (42) is true. In this case the singular point is a

branching point too. If $\alpha \geq 1$, both $\mathbf{J}$ and $\mathbf{H}$ will be indefinite. In this case, the convergence of the singular point depends on the covariance matrix of the mutants (*Leimar, 2009*).

INDIVIDUAL-BASED SIMULATIONS

Our individual-based numerical models were designed to be as similar to their analytical counterparts as possible; they differ only in being somewhat closer to natural systems. We present our methods and results using one arbitrary but illustrative parameter set concerning the size of the groups, the population and the benefit function, but we note here that the presented results are robust against variation of the parameters.

As in the analytic models, each individual is characterised by the strategy pair $x = \{q, c_A\}$. In each round, $N = 7$ individuals were sampled randomly from the well-mixed population of 1000 individuals to face the collective task. In each generation, 1000 such groups were selected, so that individual fitnesses were updated on average 7 times in each generation. The group's success in solving the task depended on the ratio of members choosing each subtask in a way that the maximal benefit was achieved when this ratio was closest to 1. The actual benefit of the group was defined by the function $\beta(i) = \frac{30}{\sigma\sqrt{2\pi}}e^{-\frac{(i-\mu)^2}{(2\sigma)^2}}$ with $\sigma = 5$, $\mu = 3.5$ and i denoting the number of group members playing one strategy. Varying σ, the behaviour of the models changes in a not monotonous, but still only quantitative manner, therefore we could use an arbitrary value for this parameter. The benefit defined by the above function was shared between group members as defined in the analytical models (I) and (II). In each case, payoffs were added to the group members' fitness, and at the end fitness was divided by the number of games an individual participated in. Those who did not participate continued to have zero fitness. After the 1000 fitness updates, 100 bouts of reproduction took place; we compared the fitnesses of two randomly selected individuals, and the offspring of the one with the higher fitness replaced the other. Every offspring inherited its parents' traits, but with probability 0.1, mutations occured in both traits of the offspring. The mutant trait values followed a uniform distribution around the parents' trait values, so they became $q' \in [q - \epsilon, q + \epsilon]$ and $c'_A \in [\sqrt[\alpha]{c_A^\alpha - \epsilon}, \sqrt[\alpha]{c_A^\alpha + \epsilon}]$ with $\epsilon = 0.005$ (and with the obvious limitations that $0 \leq q' \leq 1$ and $0 \leq c'_A \leq k$).

Branching

For testing the predictions of model (I) and (II) regarding the branching points, we designed simulations to map the parameter space. We ran 20 repeats for each parameter-set and looked for the frequency of evolutionary branching. At the beginning of a repeat the whole population was uniform with $q_0 = (1 - q_0) = 0.5$ and $c_0 = \sqrt[\alpha]{k^\alpha - c_0^\alpha} = \frac{k}{\sqrt[\alpha]{2}}$ to make sure that the trait values were in the singular point $x^* = \{\frac{1}{2}, \frac{k}{\sqrt[\alpha]{2}}\}$. In each simulation, we followed the population for a previously estimated time, enough for branching in most cases: 1.5×10^5 or 10^5 generations in model (I) and (II), respectively.

At the end of the simulations there were three types of results: (1) the population was nearly uniform, evolving to $x^{(0)}$ or $x^{(1)}$, (2) branching occurred, half of the population became specialized in subtask A, and the other half in subtask B, and (3) the population
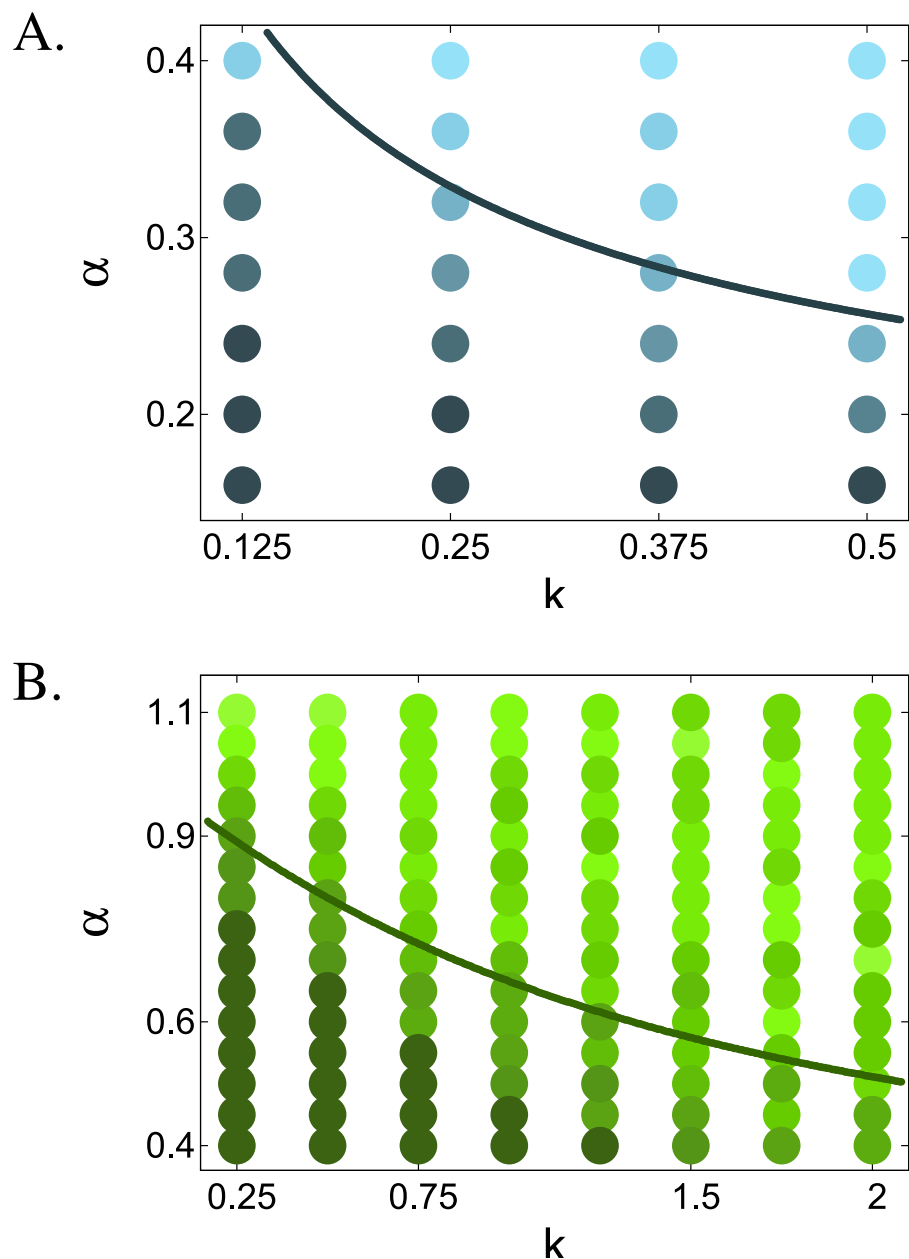
Figure 3 Frequency of branching at different _k_ and _α_ values in model (I) (A) and (II) (B). The solid curves are the predicted border lines, above which branching never occurs in the related analytical models. The darkness of the points denotes the number of branching events out of 20 at a given parameter combination with the darkest colour denoting 20 and the lightest denoting 0 branching events.

was still distributed around the original states, probably meaning that there was not enough time for branching.

After mapping a huge part of the parameter space we concluded that the simulated results (Fig. 3) match the three most important predictions of the analysis; furthermore, they reveal a new charachteristic of the stochastic model. First, when branching occured, it happened in or very close to the singular point (Fig. 4B). Second, the frequency of
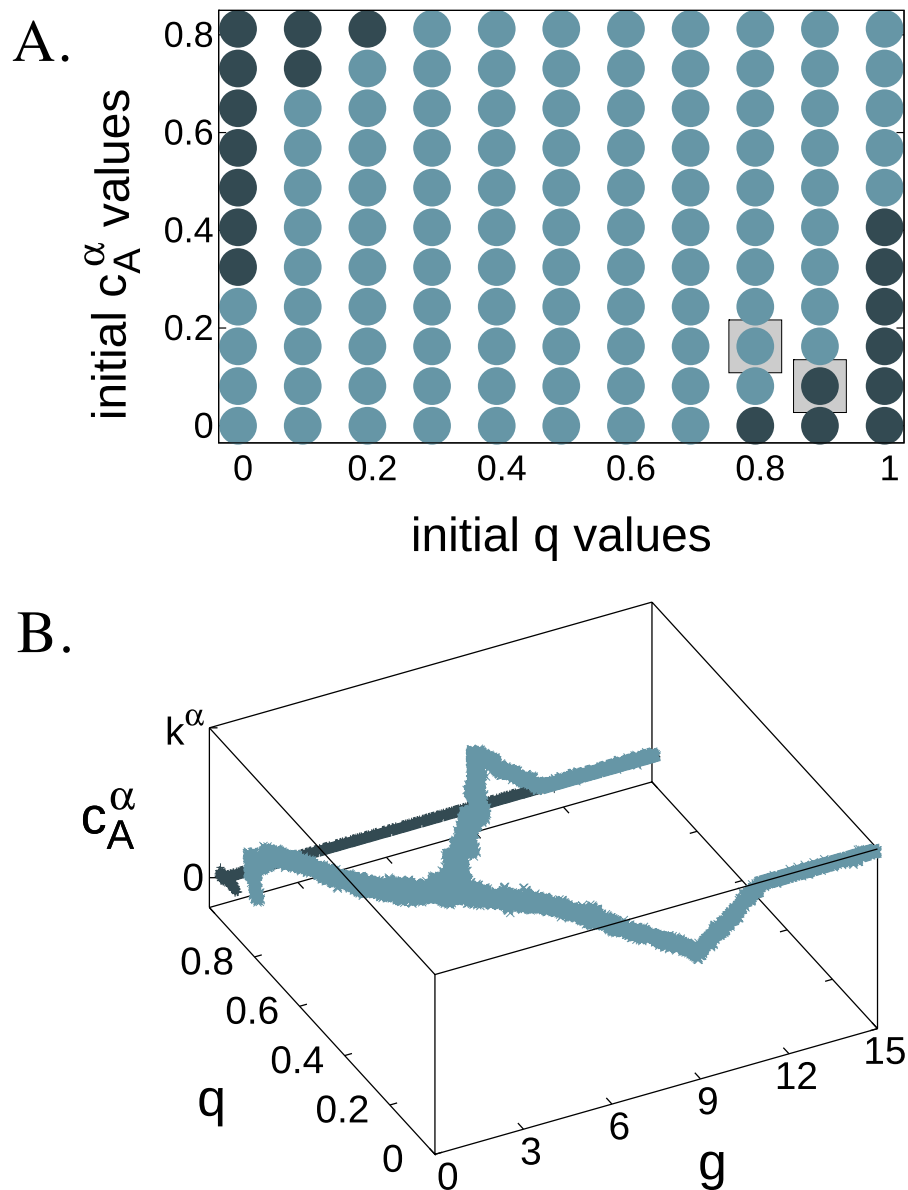
Figure 4 The effect of the initial trait values on branching in model (I). Plot A. shows the intial parameter combinations with $k = 0.25$ and $\alpha = 0.15$. Darker points denote events, where branching did not occur, and lighter points denote branching events. Plot B. shows two runs of $g \times 10^4$ generations, those denoted by squares on plot A. with respective colours.

branching followed qualitatively the analytically calculated curves (42) in both models. Third, in model (I) specialisation occured in a much narrower area of the parameter space than in model (II), as predicted in (40). We also have a fourth observation, that is, in model (II), above the predicted border line, the closer α is to 1, i.e., the weaker the trade-off is, the more branching occurs, probably as a result of the strong negative frequency dependent selection. Furthermore, while the invasion analysis is valid only in the close neighborhood of the singular point, or close to the $x^{(0)}$, $x^{(1)}$ points, the numerical simulations show that

local analysis successfully forecasts global behaviour. On the other hand, branching can occur above the critical α values as well, although with significantly smaller frequency than below the critical α (Fig. 3B), because the x^* singular point remains a branching point for every α. Thus if the selection and the stochasticity of the system allows x to move into the vicinity of x^*, occasional branching can occur. Naturally this depends on the initial conditions as well.

Finally, to see how much the benefit function counts in determining the branching behaviour, we designed some simulations with different asymmetrical benefit functions using parameter sets where branching was typical before. During these simulations we have seen that small asymmetries result in specialisation, as before, with subpopulation sizes resembling the amount of asymmetry, but large asymmetries result in homogeneous populations with only one pure strategy. The latter is probably a result of stochasticity and the singular point being too close to the unstable fixpoints x^*u or x_u^{**}.

Initial values and convergence

After testing the existence of the predicted singular point $x^* = \{\frac{1}{2}, \frac{k}{\sqrt[\alpha]{2}}\}$, we studied the effect of the initial population trait values, $\{q_0, c_0\}$, on branching in model (I). As we have shown in 'Invasion of a pure strategy by a mixed one', the $x^{(0)}, x^{(1)}$ end points can be stable against the invasion of mutants. Using the Gaussian benefit function defined above with not too small (> 0.001) k values it is easy to see that $[\beta(1) - \beta(0)]/N - k < 0$ and $[\beta(N-1) - \beta(N)]/N + k > 0$. Thus, there are additional singular points, x_u^* and x_u^{**}, as assumed above. We pointed out that the $x^* = 1/2$ singular point and the $x^{(0)}, x^{(1)}$ end points are convergence stable points, while x_u^* and x_u^{**} singular points are not. These additional singular points separate the phase-space and make the dynamics converging either to x^* or to one of the two end points.

To test these predictions, we ran simulations from 11×11 different starting points using the formerly tested parameter set ($k = 0.25$, $\alpha = 0.15$) that enabled branching. The results (Fig. 4) show us that the $x^{(0)}, x^{(1)}$ end points are indeed convergence stable such as the singular point x^*, and that there are unstable points separating the dynamics. Although our numerically estimated results do not exactly match the predicted $x_u^* = \{q = 0.0154, c_A^\alpha = 0.168\}$ and $x_u^{**} = \{q = 0.984, c_A^\alpha = 0.644\}$ points, they are quite close. We note here that in model (II) in cases when $x^{(0)}, x^{(1)}$ endpoints are unstable, there can be other singular points apart from x^*. These additional singular points do not modify our general result, that is the behaviour of x^* and can be observed only at very specific parameters.

DISCUSSION

We have highlighted how the social division of labour can lead to specialisation in a socially connected population. Our results not only tell us that such a mechanism can work, they also reveal the key contributing factors, such as the strength of the trade-off, the cost-benefit ratio, the type of the benefit sharing, and the difficulty of invading the state where every individual resolves to the same one of the two subtasks. Although most

of these constraints of specialisation have been described by others (*Wolf et al., 2007*; *Gavrilets, 2010*; *Goldsby et al., 2012*; *Rueffler, Hermisson & Wagner, 2012*; *Nakahashi & Feldman, 2014*), we believe that the social division of labour is an important new context.

In this paper, we constructed a model framework for the emergence of consistent, heritable behavioural variation via social division of labour. We have shown analytically that if population members frequently face a collective task consisting of costly subtasks, where subtask efficiencies and the probability of choosing one of the subtasks are (genetically or culturally) heritable traits and group benefit is a function of subtask frequency, then a weak trade-off between subtask efficiencies ($\alpha < 1$) can drive the population into specialisation. The specialised population members thus have optimal fitness, as DL enables them to get the highest benefit with the lowest cost, this way maintaining diversity. This occurs more easily if benefit sharing is based on the quality of contribution; that is, if those who represent a scarcer strategy, get a higher share (model II). We emphasize here that according to the general theoretical results (*Leimar, 2009*) branching is not impossible even if the trade-off function is concave ($\alpha > 1$, trade-off is strong), but then branching depends on the covariance of mutations in q and c traits. We also verified our results and extended them by numerical simulations, where more realistic conditions, such as a finite population and stochastic events were met. We show that branching is possible in the individual based model even if the singular point x^* is not convergent stable. We note here that although it would be biologically more appropriate to consider traits to be determined by many independent genes, the possibility of hybridization would complicate the outcome considerably. (There is no other difference between the two kinds of modelling, see (*Taylor & Day, 1997*)).

We have seen that one of the biggest hindrance of branching is the population being in the locally stable state of performing only one of the subtasks, that is, the invasion of the newly emerging strategy. Slightly different mutant mixed strategies cannot invade a population like this. However, in natural systems, mutations are not the only cause of behavioural change, since behaviour is shaped by environmental factors too. Our models do not include developmental plasticity, adaptations of a lifetime, environmental effects or learning but it is important not to forget about all these processes enabling behavioural shifts.

Our model describes two different types of social organisation differentiated by the way of benefit sharing. In social species, sharing a benefit collectively acquired is not rare: social spiders share a common web and their food captured together (*Grinsted et al., 2013*); several Hymenoptera species' queens share the nest/hive founded together (*Duarte et al., 2011*), and lions or chimpanzees share their prey hunted together (*Packer, Scheel & Pusey, 1990*; *Newton-Fisher, 2014*). However, we have little knowledge of a sharing system where the members' proportions are in relation to their specific value to the group. There is only one species where both empirical and experimental evidence shows that the sharing system often resembles that of model (II), and that species is us: humans (*Leuven, Oosterbeek & Van Ophem, 2004*; *Kanngiesser & Warneken, 2012*). Obviously, this is not the only reason why we find far more examples of social DL in human social systems than we find

elsewhere, but it is still in concordance with the fact that DL is more easily performed. That is, both invasion of a mixed strategy and branching appeared more easily in model (II) than in model (I). Parallel to this observation we can also see that task specialisation without sexual DL is also much more frequent in human societies than elsewhere. Our hypothesis is that the system of an inherited, life-long task specialisation could easily contribute to the appearance of different behavioural syndromes, especially in humans.

In the literature there are several hypotheses about the emergence of personality and behavioural syndromes, in some aspects similar to our idea, and we are convinced that these process are not mutually exclusive. On the contrary, they are probably working together. One significant idea is social or individual niche specialisation (*Bergmüller & Taborsky, 2010*; *Montiglio, Ferrari & Réale, 2013*), which says that each individual occupies a niche, and the more diverse a social system is, the more and more specific social niches the society can maintain. At first sight, these two ideas, social niches and social DL, seem similar, but the basic situation is quite different: social niches are needed for weakening social conflict, i.e., conflict of interest, while social DL is needed for situations with mutual interest.

Our work has emphasised that not only conflicts, but cooperation can also lead to personality diversification. We find this idea to fit well with a history of human evolution where a positive feedback loop of social DL and behavioural and skill differentiation could have contributed to the current diversity of human personalities.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Grant Disclosures

Competing Interests

The authors declare they have no competing interests.

Author Contributions

- Zsóka Vásárhelyi, Géza Meszéna and István Scheuring conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

REFERENCES

Barta Z, Székely T, Liker A, Harrison F. 2014. Social role specialization promotes cooperation between parents. *The American Naturalist* **183(6)**:747–761 DOI 10.1086/676014.

Bergmüller R, Taborsky M. 2010. Animal personality due to social niche specialisation. *Trends in Ecology & Evolution* **25(9)**:504–511 DOI 10.1016/j.tree.2010.06.012.

Boesch C. 2002. Cooperative hunting roles among Tai' chimpanzees. *Human Nature* **13**:22–46 DOI 10.1007/s12110-002-1013-6.

Boza G, Szilágyi A, Kun Á, Santos M, Szathmáry E. 2014. Evolution of the division of labor between genes and enzymes in the rna world. *PLoS Computational Biology* **10(12)**:e1003936 DOI 10.1371/journal.pcbi.1003936.

Buech. 1995. Sex differences in behavior of beavers living in near-boreal lake habitat. *Canadian Journal of Zoology* **73**:2133–2143 DOI 10.1139/z95-250.

Cartwright P. 2003. Developmental insights into the origin of complex colonial hydrozoans. *Integrative and Comparative Biology* **43**:82–86 DOI 10.1093/icb/43.1.82.

Clutton-Brock TH, Russell AF, Sharpe LL. 2004. Behavioural tactics of breeders in cooperative meerkats. *Animal Behaviour* **68**:1029–1040 DOI 10.1016/j.anbehav.2003.10.024.

Cordero OX, Polz MF. 2014. Explaining microbial genomic diversity in light of evolutionary ecology. *Nature Reviews Microbiology* **12**:263–273 DOI 10.1038/nrmicro3218.

Dall SRX, Houston AI, McNamara JM. 2004. The behavioural ecology of personality: consistent individual differences from an adaptive perspective. *Ecology Letters* **7**:734–739 DOI 10.1111/j.1461-0248.2004.00618.x.

Dingemanse NJ, Both C, Drent PJ, Tinbergen JM. 2004. Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of The Royal Society B* **271(1541)**:847–852 DOI 10.1098/rspb.2004.2680.

Duarte A, Weissing FJ, Pen I, Keller L. 2011. An evolutionary perspective on self-organized division of labor in social insects. *Annual Review of Ecology, Evolution, and Systematics* **42**:91–110 DOI 10.1146/annurev-ecolsys-102710-145017.

Gavrilets S. 2010. Rapid transition towards the division of labor via evolution of developmental plasticity. *PLoS Computational Biology* **6(6)**:e1000805 DOI 10.1371/journal.pcbi.1000805.

Geritz SAH. 2005. Resident-invader dynamics and the coexistence of similar strategies. *Journal of Mathematical Biology* **50**:67–82 DOI 10.1007/s00285-004-0280-8.

Geritz SAH, Gyllenberg M, Jacobs FJA, Parvinen K. 2002. Invasion dynamics and attractor inheritance. *Journal of Mathematical Biology* **44**:548–560 DOI 10.1007/s002850100136.

Geritz SAH, Metz JA, Kisdi E, Meszéna G. 1997. Dynamics of adaptation and evolutionary branching. *Physical Review Letters* **78**:2024–2027 DOI 10.1103/PhysRevLett.78.2024.

Goldsby HJ, Dornhaus A, Kerr B, Ofria C. 2012. Task-switching costs promote the evolution of division of labor and shifts in individuality. *Proceedings of the National Academy of Sciences of the United States of America* **109**:13686–13691 DOI 10.1073/pnas.1202233109.

Gowdy J, Krall L. 2013. The ultrasocial origin of the anthropocene. *Ecological Economics* **95**:137–147 DOI 10.1016/j.ecolecon.2013.08.006.

Grinsted L, Pruitt JN, Settepani V, Bilde T. 2013. Individual personalities shape task differentiation in a social spider. *Proceedings of the Royal Society B* **280 (1767)**:20131407 DOI 10.1098/rspb.2013.1407.

Gurven M, Winking J, Kaplan H, von Rueden C, McAllister L. 2009. A bioeconomic approach to marriage and the sexual division of labor. *Human Nature* **20**:151–183 DOI 10.1007/s12110-009-9062-8.

Hofbauer J, Sigmund K. 1998. *Evolutionary games and population dynamics.* Cambridge: Cambridge University Press.

Kanngiesser P, Warneken F. 2012. Young children consider merit when sharing resources with others. *PLoS ONE* **7**(**8**):e43979 DOI 10.1371/journal.pone.0043979.

Leimar O. 2005. The evolution of phenotypic polymorphism: randomized strategies versus evolutionary branching. *The American Naturalist* **165**(**6**):669–681 DOI 10.1086/429566.

Leimar O. 2009. Multidimensional convergence stability. *Evolutionary Ecology Research* **11**:191–208.

Leuven E, Oosterbeek H, Van Ophem H. 2004. Explaining international differences in male skill wage differentials by differences in demand and supply of skill. *The Economic Journal* **114**(**495**):466–486 DOI 10.1111/j.1468-0297.2004.00217.x.

Lloyd MM, Poulin R. 2014. Geographic variation in caste ratio of trematode colonies with a division of labour reflect local adaptation. *Parasitology Research* **113**:2593–2602 DOI 10.1007/s00436-014-3913-x.

Maynard Smith J, Szathmáry E. 1995. *The major transitions in evolution.* Oxford: Oxford University Press.

Meszéna G, Gyllenberg M, Jacobs FJ, Metz JAJ. 2005. Link between population dynamics and dynamics of darwinian evolution. *Physical Review Letters* **95**:078105 DOI 10.1103/PhysRevLett.95.078105.

Meszéna G, Kisdi É, Dieckmann U, Geritz SAH, Metz JAJ. 2001. Evolutionary optimisation models and matrix games in the unified perspective of adaptive dynamics. *Selection* **2**(**1–2**):193–210 DOI 10.1556/Select.2.2001.1-2.14.

Michod RE, Viossat Y, Solari CA, Hurand M, Nedelcu AM. 2006. Life-history evolution and the origin of multicellularity. *Journal of Theoretical Biology* **239**:257–272 DOI 10.1016/j.jtbi.2005.08.043.

Montiglio P-O, Ferrari C, Réale D. 2013. Social niche specialization under constraints: personality, social interactions and environmental heterogeneity. *Philosophical Transactions of the Royal Society B* **368**(**1618**):20120343 DOI 10.1098/rstb.2012.0343.

Nakahashi W, Feldman MW. 2014. Evolution of division of labor: emergence of different activities among group members. *Journal of Theoretical Biology* **348**:65–79 DOI 10.1016/j.jtbi.2014.01.027.

Nettle D. 2006. The evolution of personality variation in humans and other animals. *American Psychologist* **61**(**6**):622–631 DOI 10.1037/0003-066X.61.6.622.

Newton-Fisher NE. 2014. *Handbook of paleoanthropology*, *The hunting behavior and carnivory of wild chimpanzees.* Berlin: Springer Verlag, 1295–1320 (chapter).

Packer C, Scheel D, Pusey AE. 1990. Why lions form groups: food is not enough. *American Naturalist* **136**:1–19 DOI 10.1086/285079.

Penke L, Denissen JJA, Miller GF. 2007. The evolutionary genetics of personality. *European Journal of Personality* **21**:549–587 DOI 10.1002/per.629.

Rueffler C, Hermisson J, Wagner GP. 2012. Evolution of functional specialization and division of labor. *PNAS* **109**(**6**):326–335 DOI 10.1073/pnas.1110521109.

Sherman PW, Lacey EA, Reeve HK, Keller L. 1995. The eusociality continuum. *Behavioral Ecology* **6**(**1**):102–108 DOI 10.1093/beheco/6.1.102.

Simpson C. 2011. The evolutionary history of division of labour. *Proceedings of the Royal Society B* **279**(**1726**):116–121 DOI 10.1098/rspb.2011.0766.

Stampsa JA, Briffa M, Biro PA. 2012. Unpredictable animals: individual differences in intraindividual variability (iiv). *Animal Behaviour* **83**:1325–1334 DOI 10.1016/j.anbehav.2012.02.017.

Stander PE. 1992. Hunting in lions: the role of the individual. *Behavioral Ecology and Sociobiology* **29**:445–454 DOI 10.1007/BF00170175.

Taylor P, Day T. 1997. Evolutionary stability under the replicator and the gradient dynamics. *Evolutionary Ecology* **11**:579–590 DOI 10.1007/s10682-997-1513-2.

Wolf M, Van Doorn GS, Leimar O, Weissing FJ. 2007. Life-history trade-offs favour the evolution of animal personalities. *Nature* **447**:581–584 DOI 10.1038/nature05835.

Wright CM, Holbrook CT, Pruitt JN. 2014. Animal personality aligns task specialization and task proficiency in a spider society. *Proceedings of the National Academy of Sciences of the United States of America* **111(26)**:9533–9537 DOI 10.1073/pnas.1400850111.