THE EVOLUTION OF COOPERATION AND DIVISION OF LABOUR IN STRUCTURED POPULATIONS – THREE INDIVIDUAL BASED MODELS

Doctoral thesis

Zsóka Anna Vásárhelyi

Supervisor: Dr. István Scheuring

scientific advisor

Eötvös Loránd University, Faculty of Science

Doctorate School in Biology

Head of the school: prof. Anna Erdei

Theoretical and Evolutionary Biology Ph.D. Program

Head of the program: prof. Eörs Szathmáry

Department of Plant Systematics, Ecology and Theoretical Biology

2018
# Contents

1 Introduction 3

2 Cooperation in the Volunteer’s Dilemma Game 11

3 Division of labour in a collective task 31

4 Food production and behavioural specialisation 45

5 Conclusions 73

Bibliography 80

Appendix A Abbreviations and variables 97

Appendix B Supplementary for Chapter 2 100

Appendix C The analytic model of Chapter 3 106

Appendix D Supplementary for Chapter 4 116

Summary 122

Összefoglalás 123
Chapter 1

Introduction

Major Evolutionary Transitions are often characterised by cooperation and division of labour on a given level of organisation which leads to the emergence of a higher level of evolutionary unit (Maynard Smith and Szathmáry 1995, Szathmáry 2015). Thus, understanding the evolution of cooperation and division of labour is of exceptional importance. It is especially so in the case of humans, where both cooperation and division of labour operates on numerous levels and on a remarkably large scale. But cooperation of this scale is not only a result of our singular evolutionary history, but presumably it is also among the primary causes of the unusual success of the Homo sapiens (see e.g. Bingham (1999) or Hill et al. (2009)). As all this was realised many decades ago, an immense body of literature accumulated in this research field. To give an idea about the wide range of the literature, on March 08 2018 the Web of Science found 24 531 titles including the term “cooperation”, and 2287 titles including “division of labour”. Yet, however extensive this literature is, it still cannot cover all the appearing questions in connection with such a complex problem. In my thesis I attempt to answer a number of untackled questions regarding the emergence and evolutionary role of cooperation and division of labour in structured populations, with special emphasis on human social groups.

During this thesis I will discuss three studies investigating different, but related problems, as follows (for more specific questions see the corresponding chapters):

1. How can cooperation invade variously viscous populations in characteristically different social situations?
2. How does division of labour emerge during cooperative task solving in social groups, and what are the evolutionary consequences?

3. How does food production and sedentism effect the appearance of division of labour and specialisation in human social groups?

In each study I discuss an individual based model with interacting social groups. First I will shortly outline the background and my motivations behind the studies, and also how the three of them follow each other in a line of complexity. Broader model specific introductions will follow later.

**Cooperation in social dilemmas**

When studying the evolution of cooperation, two main questions arise: (i) whether cooperation can be evolutionarily stable, and (ii) whether it can spread when appearing as a rare mutant strategy (Axelrod and Hamilton 1981). For explaining the evolution of cooperation we have to find the answers to both of these questions. Unfortunately, many studies forget to answer the second one, which often is the more difficult (see e.g. Nowak and May (1992), Hauert and Doebeli (2004)).

In the first study I ask how cooperators, as rare mutants, can invade a population of cheaters in characteristically different social dilemmas, that is, situations in which collective and private interest conflict with each other. Social dilemmas are frequently experienced in nature and human societies. Well-known examples include cooperation of unicellular organisms, hunting in groups, tax paying, the open source software movement, etc. (Chuang et al. 2009, Johnson 2002, Kollock 1998, MacLean et al. 2010, Packer et al. 1990, Wilson 2011).

In the literature of theoretical work treating social dilemmas, there is a weakening but still apparent dominance of 2-person game theoretical models with two strategies, cooperation and defection (cheating). However, in most social situations there are simultaneously many more than two players active. To give just a couple of examples, the simplest colonial Hydrozoa already contain a large number of individual polyps (Cartwright 2003), social spider colonies range between one and tens of thousands (Avilés 1997), and African
wild dogs hunt in groups of 3–20 (Creel and Creel 1995). Therefore, during this thesis I only consider $N$-person problems with $N > 2$.

I compare two characteristically different $N$-person social dilemmas, the most frequently studied $N$-person Prisoner’s Dilemma (Hamburger 1973) and the less well-known, but biologically more adequate Volunteer’s Dilemma games (Archetti and Scheuring 2012). The difference between these two games is how the number of cooperators affects the benefit they provide their group with (for more details see Chapter 2). In the $N$-person Prisoner’s Dilemma Game cooperation disappears unless there is positive assortment or relatedness between cooperators (Nowak 2006). However, even if coexistence is possible, the problem of how the rare mutant strategy can spread remains a problem. In the Volunteer’s Dilemma Game cooperation is possible even in infinite well-mixed populations (Motro 1991), but the possibility of invasion remains a problem here, too. Thus, in this first study I consider the invasion of cooperators in $N$-person social dilemmas with different benefit functions, in variously viscous populations.

Division of labour, specialisation and personality diversification

As a next point, I have to give a very brief outline of my hypothesis, the Agricultural Trigger Hypothesis (currently in manuscript form (Vásárhelyi 2018)) which gives the theoretical background and motivation behind the second and the third studies. The Agricultural Trigger Hypothesis proposes that the Neolithic transition induced a set of feedback loops resulting in a greater diversity of the human behaviour and personality.

According to the hypothesis, the parallel appearance of a number of life history and subsistence changes, characteristic of the Neolithic transition, together contributed to human personality diversification through the emergence of social division of labour (see in the next section) and specialisation. Among these changes were the appearance of food production, sedentary life and greater group and population sizes (Banning 2012, Maher et al. 2012, Simmons 2012, Gowdy and Krall 2013). Obviously, this pattern of changes already contains a set of positive feedback loops. For example, while large scale food storage is impossible without some kind of sedentism, sedentary life is easier if there is food storage, because local resources usually fluctuate (Binford 1980, Hamilton et al.
CHAPTER 1. INTRODUCTION

Thus, sedentism and food storage are tied in a positive feedback loop. However, what is less obvious is the emergence and the role of specialisation and social division of labour (SDL), that is, permanent division of labour within age and sex groups (Vásárhelyi et al. 2015). Studying the extensive literature on recent hunter-gatherers it is astonishing that I was unable to find any account on either permanent SDL or task specialisation between hunter-gatherer men or between women. What I did find is more the lack of such DL. E.g. Marlowe (2010) writes about the Hadza that “...there is no role specialization of any kind other than the sexual division of labor. [...] There is simply no routine division of labor except that based on sex.” What I did find are only accounts on moderate amounts of specialisation. E.g. in the Tsimane, the Ache or the Kubo tribes, while there is no kind of general same-sex task specialisation, hunters often specialise in different weapons or target species (Dwyer and Minnegal 1993, Gurven and Von Rueden 2006). And indeed, leading scientists also confirmed that in the literature there is no systematic data on hunter-gatherer specialisation (Kim Hill, personal communication), and that SDL is virtually absent in recent mobile hunter-gatherer groups (Frank Marlowe, personal communication). Based on this we can also assume that both SDL and specialisation were absent or very limited in most Palaeolithic groups, too. This notion is supported by that the first clear evidence of specialised individuals comes only from the Pre-Pottery Neolithic B (Finlayson and Warren 2012). Thus we can ask two questions: (i) What circumstances are needed for SDL and specialisation to appear? (ii) What are the consequences of SDL and specialisation with regard to behaviour and personality? The second study considers the first question, and the third study tackles both of these questions. Naturally, modelling such processes in their full complexity is practically impossible, thus my models are also simplified.

Division of labour

As I have mentioned before, in the second and third studies I turn my attention towards the problem of division of labour (henceforth DL) in cooperating groups. In the following, since the theoretical and empirical literature on division of labour is so diverse, I find it useful to give a short outline and classification of the phenomena.
DL, the cooperative separation of different tasks, can be observed on every level of biological organisation from genes to social groups. However, the literature is somewhat confusing in what DL actually is, as authors put the same label on very distinct phenomena or mechanisms. Therefore, I categorise DL along two dimensions: whether it is linked to or unlinked to reproduction, and whether it occurs within an individual or between individuals (Table 1.1). 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), such as the sterile castes in eusocial insect societies or the somatic cells after the germ-soma differentiation in multicellular organisms. Reproductive DL in a broader sense happens between other-sex individuals, too. Thus, reproductive DL could allow reproduction-linked DL to arise. Such an example could be the functional and physical specialisation of reproductively inactive female ants and Trematoda larvae to defence (Duarte et al. 2011, Lloyd and Poulin 2014). 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 are therefore not the direct results of reproductive asymmetries: I call this reproduction-unlinked DL. Examples include complementary enzymes in a protocell (Boza et al. 2014), the specialisation of limbs or organs (Rueffler et al. 2012), task-specialisation in social spider colonies (Wright et al. 2014), and the cooperative hunting of lions and chimpanzees (Stander 1992, Newton-Fisher 2014).

As the above mentioned 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, I differentiate between within and between individuals (Table 1.1), without getting into the question of what actually an individual is (see e.g. Janzen (1977)). Obviously, boundaries between the categories outlined above and in Table 1.1 are sometimes blurred and controversial, yet I think that such a division still helps to see differences and similarities between the densely connected phenomena.

Each of the four categories circumscribes a type of DL that has been studied on its own (for examples and references see Table 1.1), but one can also unfold a bigger picture with several interrelations between categories. For instance, the process by which between indi-
<table>
<thead>
<tr>
<th>Within individuals</th>
<th>Reproduction-linked DL</th>
<th>Reproduction-unlinked DL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Germ-soma differentiation in <em>Volvox</em> colonies (Michod et al. 2006)</td>
<td>Paralogous genes (Rueffler et al. 2012)</td>
</tr>
<tr>
<td></td>
<td>Non-reproductive polyps in Hydrozoa colonies (Cartwright 2003)</td>
<td>Complementary enzymes in a protocell (Boza et al. 2014)</td>
</tr>
<tr>
<td></td>
<td>Sterile workers in Hymenoptera societies (Sherman et al. 1995)</td>
<td>Specialised limbs or organs (Rueffler et al. 2012)</td>
</tr>
<tr>
<td></td>
<td>Sterile Trematoda rediae (Lloyd and Poulin 2014)</td>
<td>Cooperative hunting in chimpanzees (Boesch 2002)</td>
</tr>
<tr>
<td></td>
<td>Task specialisation due to the reproductive hierarchy in meerkats (Clutton-Brock et al. 2004)</td>
<td>DL between reproductive female spiders in a common web (Grinsted et al. 2013, Wright et al. 2014)</td>
</tr>
<tr>
<td></td>
<td>Biparental care in different animal societies (Barta et al. 2014)</td>
<td>Task specialisation of sterile Hymenoptera workers (Duarte et al. 2011)</td>
</tr>
<tr>
<td></td>
<td>Sexual DL in human societies (Gurven et al. 2009)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1: A possible classification of the diverse phenomena labelled as division of labour (DL). DL can occur within or between individuals and can be linked or unlinked to reproduction, that is, independent of it (for further details see the main text). This thesis concentrates on the section with grey background that I call social DL.
individual DL becomes within individual DL is where a higher level or organisation emerges from a cooperative group of lower level organisms (or organisations). This mechanism, as I have already mentioned before, often contributes to major transitions in evolution. 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 and 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 could be put in both rows, that is, in all elements of Table 1.1.

In the second and third studies I am mostly interested in permanent reproduction-unlinked DL between individuals, which I call Social DL (SDL). There are intriguing examples of non-human SDL in the animal kingdom. The reproductive and social DL of some Hymenoptera we already have touched. Here, the behaviour and the morphological features of workers are primarily regulated by the environment (through e.g. larval nutritional factors or the queen’s hormones and pheromones), but gene-environment interactions also play a role (Anderson et al. 2008). Colonial spiders provide another example for SDL. In some spider species numerous individuals live in a common web, and solve cooperatively a number of tasks, such as offspring provisioning, web maintenance, prey capture or colony defence (Lubin and Bilde 2007). But spiders not only cooperate, they also divide labour. Moreover, in certain species, different behavioural types within a colony typically solve different tasks with different probabilities (Grinsted et al. 2013, Wright et al. 2014). Importantly, in these species reproduction is not a privilege of one or a restricted group of females (Darchen and Delage-Darchen 1986). Unfortunately, not much is known yet about the regulation of behaviour in these cases, nor about the selective and ecological circumstances behind this SDL.

We have seen examples for behavioural specialisation and SDL in species with relatively limited cognitive abilities. Presumably, the more complex an animal’s cognitive
abilities, the more diverse its behaviour. Therefore, moving towards more intelligent birds or mammals, we see a greater variability, a greater number of interaction types between individuals, and also a greater number of regulatory mechanisms acting on behaviour. These latter include internal factors, such as hormone levels (Trumble et al. 2015) or different receptor variants (Takahashi and Miczek 2013, Hawi et al. 2015), and external factors, such as resource availability (Boccia et al. 1988, Sima et al. 2016) or socioeconomic status (Van de Mheen et al. 1998, Jonassaint et al. 2011). It is surprising then that, despite the great behavioural plasticity and high levels of sociality, apart from humans, there are no systematic data on permanent SDL in animals with advanced cognitive abilities. Nevertheless, it is not only common, but even inevitable in post-agricultural humans. As we will see later, this observation motivated the third model. Note also that in the second and third studies I assume that a social group consists of individuals who all contribute more or less to the common good, that is, I do not consider the problem of free riding any more.

To summarise, the first study, Chapter 2, investigates the role of the benefit function and the possibilities of a cooperator invasion in variously viscous populations of cheaters. It focuses on the Volunteer’s Dilemma Game, which is a biologically relevant model of many social dilemmas, and compares it with the N-person Prisoner’s Dilemma Game. This chapter is based on Vásárhelyi and Scheuring (2013). The second study, Chapter 3, asks what circumstances are needed for SDL to appear in a social group, and what potential consequences arise after it appeared. It discusses an analytically tractable individual based model of SDL. This chapter is based on Vásárhelyi et al. (2015), but includes additional results, too. Finally, the third study, Chapter 4, goes even further and looks for whether the human behavioural complex or personality could be affected by the appearance of SDL and specialisation. It discusses an individual based model on how the ecological environment could affect personality diversification. This chapter is based on a submitted manuscript (Vásárhelyi and Scheuring 2018). I note here that in the following I will often use plural, as the studies contain not only my, but also my colleagues’ work, too. Yet, everything I included in the main text is primarily my own work. In the following I will discuss the above mentioned studies each on their own, with no reference to the others. However, at the end of this thesis there follows a general Conclusion (Chapter 5.).
Chapter 2

Cooperation in the Volunteer’s Dilemma Game

Introduction

Situations, where a conflict of interest stands between individuals and their social groups are ubiquitous in nature. Concerning these situations, called social dilemmas, two key questions emerge: (i) What kind of mechanism maintains cooperative or altruistic behaviour in situations where selfishness seems more beneficial? And an even more challenging question: (ii) How can cooperation invade populations where selfishness is dominant?

As I have mentioned in Chapter 1, most game theoretical models of social dilemmas contain only two interacting participants. The most popular ones, the Prisoner’s Dilemma (PD) (Hamilton 1971, Axelrod and Hamilton 1981) and the Snowdrift Game (SD) (Sugden 1986, Doebeli and Hauert 2005) are simple models that capture some key attributes of the problem. Due to decades of research on the PD, it is common knowledge now that well-mixed populations are dominated by defection, but in the case of structured, spatial, or socially inhomogenous populations, the stable coexistence of cooperation and defection or even the dominance of cooperation can occur. (Nowak 2006, Nowak and May 1992, Perc and Szolnoki 2010, Szabó and Fáth 2007b). However, as I have shown before, social situations typically contain more than two individuals. Indeed, a rapidly growing body of experimental research (Chuang et al. 2009, Damore and Gore 2012, Lee et al.

Let us outline briefly the main results obtained from the simplest and most frequently used models of N-person social dilemmas. Consider an infinite well-mixed population with strategies $C$ (cooperation) and $D$ (defection), where selection is studied by replicator dynamics (Hofbauer and Sigmund 1998). In the model framework for N-person SD games (NSD), the coexistence of cooperators and defectors is the only stable fixed point of replicator dynamics. However, the frequency of cooperators decreases with $1/N$ at equilibrium (Souza et al. 2009, Zheng et al. 2007). In the N-person PD game (NPD), which builds on the 2-person PD game, cooperation disappears if $c/b > 1/N$ where $c$ is the cost and $b$ is the benefit of the cooperative act (for details see Appendix B). It is generally argued that this relation is valid for most biologically reasonable situations (Hauert et al. 2006a, Archetti and Scheuring 2012). To summarise, the frequency of cooperators is marginally low or zero in large well-mixed populations where individuals play the NSD or the NPD games in larger groups. These results sharply contradict the numerous observations of cooperation in groups of bacteria, animals, and in human communities (Kollock 1998, Lee et al. 2008, Rainey and Rainey 2003, Yip et al. 2008).

Naturally, the condition of a large well-mixed population does not necessarily hold for real situations, because, in many cases, there is some kind of positive assortment among cooperators (e.g. colonies of bacteria or groups of cooperative animals or humans, etc.). Positive assortment or relatedness, giving opportunity for kin-selection to act, enables cooperators to outcompete or coexist stably with defectors even in NPD-like situations (Hamilton 1964, 1971, Van Baalen and Rand 1998, Nowak and May 1992, Nowak 2006, Perc et al. 2013). Thus, the literature gives us the impression that positive assortment (or relatedness) is necessary for cooperation to persist in NPD/NSD-like situations. However,
in addition to positive assortment, there is another candidate mechanism that can explain how the production of public goods exists stably in well-mixed, natural systems, which is to apply an alternative and more adequate N-person game to the NSD or the NPD (Motro 1991, Hauert et al. 2006a, Archetti 2009, Pacheco et al. 2009, Archetti and Scheuring 2012). One of these alternatives is the Volunteer’s Dilemma Game (VD) (or Threshold Public Goods Game) (Diekmann 1985, Bach et al. 2006, Archetti 2009, Archetti and Scheuring 2011).

The crucial difference between the NPD and the VD games is the shape of the public goods function they use. In the NPD, the public good increases linearly with the number of cooperators, \( i \) (Fig. 2.1). However, in real biological systems, this linearity is at most a rare exception, and certainly not the rule. The well-documented case of cooperative hunting in groups is a good example, in which per capita success and benefit change non-linearly with the number of participants (Bednarz 1988, Packer et al. 1990, Stander 1992, Creel and Creel 1995, Yip et al. 2008, Hooper et al. 2015). Similarly, cooperative nesting and breeding of vertebrates (Rabenold 1984) and the fruiting body formation of social amoebas (Bonner 2008) are examples of non-linear public goods games. In microbial communities, where the public good is based on specific molecules like replication enzymes (Turner and Chao 2003), adhesive polymers produced by viruses (Rainey and Rainey 2003), antibiotic resistance in bacteria (Lee et al. 2008) or invertase enzyme in yeast (Gore et al. 2009), the effect of enzyme production, and thus the amount of obtainable common good must generally be a saturating function of the molecule concentration (Hemker and Hemker 1969, Ricard and Noet 1986, Mendes 1997, Eungdamrong and Iyengar 2004). Following this train of thought, the public goods function \( (B(i)) \) used when modelling these kind of dilemmas should be a monotonously saturating curve with an inflection point somewhere, i.e. a sigmoid one (Fig. 2.1). It is quite a challenge to analyse a model with this benefit function instead of the linear NPD. In order to simplify the problem, specific \( B(i) \) functions are defined (Motro 1991, Bach et al. 2006, Hauert et al. 2006b, Archetti 2009).

It is biologically realistic to assume that the transition phase from accelerating to discounting is steep at the inflection point, and thus, as a specific model, we can use a step function for \( B(i) \) (Archetti 2009):
Figure 2.1: Schematic picture of characteristically different public goods as a function of number of cooperators. A general saturating non-linear $B(i)$ function with inflection point at $h_{cr}$ (dashed line), the linear $N$-person Prisoner’s Dilemma game (dotted line) and the $N$-person Volunteer’s Dilemma game with threshold at $h_{cr}$ (continuous line).

\[ B(i) = \begin{cases} 
0 & \text{if } i < h_{cr} \leq N \\
b & \text{otherwise}, 
\end{cases} \]

which defines the generalised $N$-person Volunteer’s Dilemma Game (Archetti 2009, Diekmann 1985). That is, at least $h_{cr}$ cooperators (volunteers) are needed for the high benefit, otherwise cooperation does not achieve net benefit (in the classical VD game $h_{cr} = 1$). This model is more tractable, and its dynamical behaviour remains qualitatively the same as it would be for smooth but similar functions of $B(i)$ (Fig. 2.1) (Archetti and Scheuring 2012). We note here that the problem of a collective action in public goods game with non-linear benefit functions is a classical one in sociology and political science (Hardin 1982, Maxwell and Oliver 1993, Oliver 1993, Ostrom 2003, e.g.). However, studying this problem in the framework of evolutionary game theory has become an important issue only recently.

Using the replicator dynamics, one can look for the stable states of the dynamics in the VD for infinitely large well-mixed populations, that is, one can tell whether a stable
coexistence between cooperators and defectors is possible. It can be shown thus that there is a parameter range of $N$, $c$ and $b$, for which complete defection is the only stable state of the dynamics ($x^* = 0$, if $x$ denotes the frequency of cooperators, and $x^*$ is the fixed point). However, in the other range, $x^* = 0$ and $x^* = x_s < 1$ are the stable fixed points, and $x^* = x_u < x_s$ and $x^* = 1$ are the unstable fixed points of Eq. S4 (Archetti and Scheuring 2011, 2012). In other words, there is a parameter range where a stable coexistence of cooperators and defectors is possible in the VD (for details, see Appendix B). Naturally, if the population structure allows positive assortment for cooperators and the social dilemma is described by the VD, cooperators coexist with defectors or dominate structured populations even more easily (Boza and Számadó 2010, Szolnoki and Perc 2010, Perc et al. 2013).

It is important that $r$, which is a complex function of $b/c$ and $N$, can be much smaller than $N$ for the VD and for similar games with smooth but highly nonlinear benefit functions, so stable coexistence of $D$ and $C$ strategies at $x_s$ is a typical solution of these non-linear public goods games (Archetti and Scheuring 2011, 2012). We have to note here that the cooperative state is evolutionary unstable if investment (that is, $c$ and $b = rc$) is an evolvable parameter of the VD, but remains stable for smooth sigmoid benefit functions (Deng and Chu 2011). Since the strategies and consequently the investments of the presented model are fixed, this problem does not appear here. The validity of this assumption is supported by numerous microbial systems where only a few discrete types are isolated, typically a cooperator and a cheater strain (West et al. 2002, Velicer 2003, Greig and Travisano 2004, Chuang et al. 2009, Gore et al. 2009, e.g.). Since $x^* = 0$ remains an alternative stable fixed point, rare $C$ strategies never invade the population of $D$ (see question (2) above). Stable coexistence of $C$ and $D$ strategies can be attained only if the initial frequency of cooperators is higher than $x_u$ (Fig. 2.2).

The deterministic model framework assumes an infinitely large population, but in real finite populations, where births and deaths are stochastic events, invasibility has to be measured by the probability of a successful invasion (spread or fixation) of an initially rare strategy (Nowak et al. 2004, Kurokawa and Ihara 2009, Gokhale and Traulsen 2010). In this study, we are concerned with the likelihood of rare cooperators invading populations dominated by defection, that is, the second key question addressed above. To
compare different situations, we use models of cellular automata in which individuals play stochastic NPD and VD games or a general saturating non-linear $N$-person game with their neighbours, while dispersal can vary from low to very high for modelling cases ranging from highly viscous and spatially structured populations to well-mixed populations.

![Figure 2.2: Coexistence of cooperators and defectors in the analytic VD model. The fitness of cooperators (gray) and defectors (black) are depicted as a function of the frequency of cooperators. Filled circles denote stable, open circles denote unstable fixed points of the dynamics (see Eq. S4). Arrows show the direction of motion, thus $x_u$ and $x_s$ are the nontrivial unstable and the stable fixed points, respectively. $c/b = 0.2$, $h_{cr} = 4$, $N = 9$.](image)

**Models and Methods**

To study invasion probabilities we consider a toroidal square lattice grid of $L^2$ (generally 100x100) nodes with at most one individual per grid point. Every individual interacts with its eight nearest neighbours (Moore neighbourhood), thus $N_g = 9$. Individuals differ only in their strategies: cooperators ($C$) always cooperate with cost $c$ and defectors ($D$) never do. Payoffs are determined by interactions with neighbours. Thus, the fitnesses of
cooperators and defectors are

\[ W_D = W_0 + B(i) \]  \hspace{1cm} (2.1)  
\[ W_C = W_0 + B(i) - c, \]

where \( B(i) \) is the benefit function of the focal individual having \( i \) number of cooperators in the neighbourhood, and \( W_0 \) denotes the basic payoff, independent of the individual’s strategy (in most simulations, \( W_0 = 0.8 \)). The strength of selection can be estimated by the maximal possible payoff difference, that is, by \( \left[ (W_0 + b) - (W_0 - c) \right]/W_0 = (b + c)/W_0 \). In most cases, this value was \( 0.11/0.8 = 0.1375 \) (\( b = 1, c = 0.1 \)). During simulations, we examined the effect of three different benefit functions discussed below.

In case of the VD game, \( B(i) = b \) if the number of cooperators \( (i) \) is equal or higher than the critical threshold, \( h_{cr} \) in the neighbourhood. If \( i < h_{cr} \), the focal individual remains to have the basic payoff. Consequently, using the notation + for higher and ◦ for lower \( B(i) \) values, payoffs are defined by the following equations:

\[ W_D^+ = W_0 + b \]  \hspace{1cm} (if \( i \geq h_{cr} \))  
\[ W_C^+ = W_0 + b - c \]  \hspace{1cm} (2.2)  
\[ W_D^\circ = W_0 \]  \hspace{1cm} (if \( i < h_{cr} \))  
\[ W_C^\circ = W_0 - c \]

The linear benefit function (NPD game) is defined by \( B(i) = \frac{b}{9}i \), where \( B(i) \) is the value of the benefit, which depends on the number of cooperators \( i \), and the coefficient \( \frac{b}{9} \) is used here to regulate the maximal benefit to be \( b \). The sigmoid benefit function is defined by \( B(i) = b \frac{S(i) - S(0)}{S(9) - S(0)} \) where \( S(i) = \left[ (1 + e^{-\lambda(i-h_{cr})})^{-1} \right] \). Thus \( B(i) \) depends on the number of cooperators \( i \), \( h_{cr} \) is the threshold, \( \lambda \) regulates the steepness of the function at the inflection point, and the benefit varies between 0 and \( b \); that is, \( B(0) = 0, B(9) = b \). During the simulation we used \( \lambda = 1.5 \).

At the beginning, we place \( h_0 < L^2 \) cooperators randomly on the grid, while the rest of the grid points \( (L^2 - h_0) \) are occupied by defectors. First we compute the payoff of every individual with the method discussed above within one Monte Carlo cycle (MC). After this, individuals die with a constant \( \omega \) probability, independent of the strategies (in most cases \( \omega = 0.2 \)), leaving empty sites. Empty sites can be occupied again by the offspring
of living individuals in their neighbourhood. Every neighbour of an empty site \((i_o,j_o)\) produces progeny with probability \(W_f/\sum_{ijM_o} W_{ij}\) (Nowak and May 1992, Nowak et al. 1994), where \(W_f\) is the payoff of the focal neighbour of the empty site and \(\sum_{ijM_o} W_{ij}\) is the sum of the payoffs in neighbourhood \(M_o\) of \((i_o,j_o)\). Thus, neighbours place progeny to the empty site with probability proportional to their relative payoff. We note that other update rules also could be used here (Szabó and Tóke 1998), however, the benefit of the selected rule is its clear biological interpretation: progenies are in local competition, and their success is proportional to their local relative fitness. The next MC starts when all empty sites, which had at least one non-empty neighbouring site are occupied again. Although it is possible that at the beginning of a new MC there remain empty sites on the grid, that is if empty patches of at least \(3\times3\) grid points emerge after individuals die, it is realised with very low probability \((\omega^0)\) at \(\omega = 0.2\). Therefore the presence of empty sites at the beginning of a new MC has only a marginal effect on the system, if any. We did not find characteristically different results when varying \(\omega\) between 0 and 0.4, the only change appeared in the speed of the dynamics.

Using the algorithm explained above, we examined both extreme and intermediate population structures:

1. The highly viscous model, in which the population is not mixed in the simulation (apart from the weak mixing generated by reproduction to neighbouring sites). We call this the “viscous model”.

2. The well-mixed model, in which each MC cycle is followed by intense mixing, such that every individual exchanges its position with a randomly chosen other individual.

3. The variable mixing model, i.e. variable viscosity, in which each MC cycle is followed by a given number (denoted by \(m\)) of mixing steps. One step means the exchange of the position of a randomly chosen individual with its randomly chosen neighbour.

On average, \(m = L^2 = 10^4\) mixing steps are needed for each position to be chosen.

Our basic parameter set consisted of the \(h_{cr} = 3\), \(c/b = 0.1\), \(\omega = 0.2\) and \(W_0 = 0.8\) values. This seemingly arbitrary decision has relevance as long as it represents parameters
in the middle of the relevant area and can be used as a starting-point of further and extended studies. Consequently, we also examined the system with parameters \( h_{cr} = 2, 4 \) and \( c/b = 0.15, 0.2 \). When varying \( c/b \) we also varied the value of \( W_0 \) to hold the maximal payoff differences and thus the strength of selection \((b+c)/W_0\) constant across simulations. We also varied the strength of selection by using parameters \( W_0 = 0.1, 3.2 \) along with the original \( c/b = 0.1 \).

To estimate the relative abundance of cooperators in the stable polymorph population, we ran simulations with an initial frequency of cooperators \( x_0 = h_0/L^2 = 0.5 \). For \( h_{cr} \leq 7 \) the unstable equilibrium \( (x_u) \) of the analytical model is below this value, so we assume that cooperators can reach stable frequency similar to \( x_s \) from this initial value instead of being eliminated at our parameter set. Simulations were run for \( 3 \times 10^4 \) MC cycles. In each case, we counted the average ratio of cooperators from the last thousand generations, and then we counted the mean of hundred independent averages of the same parameter sets. To make comparisons, we also evaluated the stable and unstable fixed points of the analytic model (Archetti and Scheuring 2011).

In the invasion experiments, we modified the initial ratio of cooperators \( x_0 \) on a continuum from \( 1/L^2 \) to 0.4. We considered invasion to be successful when the ratio of cooperators exceeded a critical threshold fraction, \( x_{cr} \). Since in the previous experiment we observed that cooperators always remain in polymorph equilibrium with defectors (or go to fixation) if the initial ratio, \( x_0 \) was 0.5, we simply chose \( x_{cr} = 0.5 \) as the threshold for indicating successful invasion of cooperators. Obviously, invasion was considered being unsuccessful if invading cooperators were eliminated. The invasion probability \( \rho(x_0) \) is the ratio of successful experiments divided by all experiments. Although it is true that if we had let the simulations run a very long time, one of the strategies would have surely gone extinct, it can be shown that if a polymorphic equilibrium is stable in an infinite population, it remains at the corresponding metastable state for a very long time in a stochastic finite population as well. In our cases, simulation time was generally observed to be much smaller than the average fixation time from the metastable polymorphic state. We note a situation in the following section when this was not the case. To compute \( \rho(x_0) \), we used a different number of independent simulations between \( 5 \times 10^2 \) and \( 2 \times 10^5 \), depending on the parameter sets.
Finally, we examined the role of the population size, using alternative grids with sizes varying between 50x50 and 250x250 grid points. For these simulations, we used the highly viscous and the well-mixed models. In the well-mixed model with larger grid sizes ($L^2 \geq 200x200$) we have set the $x_{cr}$ critical threshold closer to the stable equilibrium (thus $x_{cr} = 0.45$ in these cases), to optimise simulation time.

Results

Using the analytic model (Archetti and Scheuring 2011) we evaluated the nontrivial fixed points of the VD and we estimated the stable equilibria in the well-mixed and viscous models defined above (Fig. 2.3). Results show that while our well-mixed and the analytic model behave very similarly, the viscous model differs notably from them: the number of cooperators in the stable equilibrium is much higher for every $h_{cr}$ than in the well-mixed models. Further, in the viscous model, the population soon reaches pure cooperation except for low values of $h_{cr}$. Not having enough time and computational capacity, we could not estimate the unstable fixed points.

The probability of invasion by the cooperator strategy was measured as a function of the initial number of cooperators. In the well-mixed case, the invasion probability $\rho(x_0)$ follows a sigmoid curve (Fig. 2.4). If the initial number of cooperators, $x_0$ is below the unstable fixed point, the invasion probability is lower than it would be in a neutral case. We consider neutrality, when there is no selection and thus every individual has equal probability of fixing, i.e. $\rho(x_0) = h_0/(L^2 x_{cr})$, where $h_0$ is the initial number of cooperators and $x_{cr}$ is the critical threshold proportion over which we considered invasion successful (Nowak et al. 2004). Varying parameters $h_{cr}$ and $c/b$ in a given parameter range, the shape of the $\rho(x_0)$ curve remains characteristically the same, only it moves from left to right as $h_{cr}$ or $c/b$ increases (Fig. 2.4). Varying the strength of selection, the $\rho(x_0)$ curve remains qualitatively the same. In the viscous model, the $\rho(x_0)$ function follows a concave saturating curve. The curve is above the neutral case in all measurements, even if $x_0 = 10^{-4}$, i.e. if there is only one cooperator in the population at the beginning (Fig. 2.4 B, D). Varying parameters $h_{cr}$, $c/b$, or the strength of selection in a given parameter range, the shape of the $\rho(x_0)$ curve remains qualitatively the same (Fig. 2.4). In the
well-mixed case, the number of cooperators fluctuates more strongly around $x_s$, than in the viscous case (Fig. 2.5). Consequently, if the polymorphic stable state is at a low cooperator concentration, the cooperator strategy gets eliminated by random drift more easily in the well-mixed populations than in the viscous ones (fixation does occur within the simulation time scale). That is why there are no data for $c/b = 0.2$ and $h_{cr} = 4$ in Fig. 2.4/A. There, cooperators always disappeared within some hundreds of generations, independent of their initial frequency.

When using smaller and larger grids we found that grid size (population size) did not alter the results qualitatively, that is, the shapes of the invasion probability curves remained the same. However, these simulations revealed that while in the well-mixed case the probability of invasion depends on the initial fraction of cooperators, in the viscous model it depends on the initial number of cooperators (Fig. 2.6). We also found that $\rho(x_0)$ of the well-mixed model converges to a step function as the population size (grid size) increases (see inset in Fig. 2.6).
Figure 2.4: Invasion probability ($\rho(x_0)$) of the cooperator strategy, as a function of the initial ratio of cooperators ($x_0$) in the well-mixed (A and C) and the viscous (B and D) models of the VD. Data differ in parameters $c/b$ (A and B) and $h_{cr}$ (C and D). Above: the triangle denotes $c/b = 0.1$, the circle denotes $c/b = 0.15$, and the square denotes $c/b = 0.2$. Below: the triangle, circle and square denote the $h_{cr} = 2$, $h_{cr} = 3$ and $h_{cr} = 4$ cases, respectively. Curves are fitted for visualisation. The straight line denotes the hypothetical neutral case, when there is no selection.

In the variable mixing model, we found that with increasing mixing, $\rho(x_0)$ decreases monotonically (Fig. 2.7). $4 \times 10^4$ mixing steps/MC returns a good approximation of the well-mixed case and, obviously, if the number of mixing steps tends to zero, we recover the viscous case. The probability of fixation decreases continuously as mixing increases,
Figure 2.5: The frequency of cooperators in the first $10^4$ generations of the well-mixed (on the right) and viscous (on the left) models of the VD. At the beginning, the fractions of cooperators and defectors are equal, $c/b = 0.1, h_{cr} = 3$, other parameters are defined in the text.

thus a high invasion probability of cooperators still occurs when mixing is rather intense, though not perfect (Fig. 2.7).

When using the linear benefit function (NPD game) we found that the invasion probability of $C$ is higher than that of a neutral strategy in the viscous case (similar to the basic model using the step function, although there the difference is more pronounced). However, invasion probability is constantly zero in the well-mixed case (Fig. 2.8). Using the sigmoid benefit function, the shape of $\rho(x_0)$ is generally much more similar to that of the step function model than to the linear model. However, the $\rho(x_0)$ values of the sigmoid function model are always lower than those of the step function model.

This gives us the impression that sigmoid functions are generally closer to the step extreme than to the linear extreme in behaviour. Naturally, an infinite number of sigmoid
CHAPTER 2. COOPERATION IN THE VOLUNTEER’S DILEMMA GAME

Figure 2.6: Invasion probability of the cooperator strategy with varying grid sizes ($L^2$) in the viscous (left) and the well-mixed (right) cases. In the viscous case, the invasion probability depends on the initial number ($h_0$) of cooperators. The triangle denotes grid size $L^2 = 100 \times 100$, the circle denotes $L^2 = 200 \times 200$, and the square denotes $L^2 = 50 \times 50$. In the well-mixed case, the invasion probability depends on the initial fraction ($x_0$) of cooperators. Darker points represent larger grids. The used and plotted grid sizes are 50x50 (lightest gray), 70x70, 100x100, 150x150, 250x250 (black). The inset figure on the right shows steepnesses of fitted curves at their inflection points ($St$) as functions of the grid size ($L^2$) on logarithmic scales. Data are obtained from simulations using the above mentioned and two other (120x120 and 200x200) grid sizes. The lines are fitted for visualisation.

Benefit functions can be defined, and their steepnesses at the inflection points determine which extremes (VD or NPD) they are closer to (Archetti and Scheuring 2012). Repeating the simulations at different grid sizes, we found that invasion probability depends on the initial number of cooperators for NPD and sigmoid benefit functions as well (results not shown).

To better understand the surprising success of the invading cooperator strategy in the viscous case, we examined in more detail what happens at the beginning of invasion if we put a single cooperator amongst defectors. We use the term “elementary state” to indicate when there are 1, 2, 3, etc. cooperators in the population. These states are denoted by $e_1$, etc.
Figure 2.7: Invasion probability ($\rho(x_0)$) as function of mixing ($m/MC$). Left: Invasion probability ($\rho(x_0)$) at varying mixing and fixed initial fractions of cooperators. Triangle denotes the $x_0 = 0.05$, circle denotes the $x_0 = 0.1$ case. Right: Invasion probability as function of the initial fraction ($x_0$) of cooperators at different levels of mixing. Triangle denotes the well-mixed case, square denotes the case with $m = 2 \times 10^4/MC$, circle denotes the one with $m = 10^4/MC$, upside-down triangle denotes the viscous simulations. Curves are fitted for visualisation, the straight line denotes the hypothetical neutral case. $c/b = 0.1, h_{cr} = 3$.

e_2, e_3, etc. During the simulations, we estimated the probabilities of transitions between these states (i.e. the probability of the number of cooperators increasing by one) as $p\{e_i \rightarrow e_{i+1}\} = p\{e_{i+1}\}$ for small $i$-s, as follows. Initially we put $i$ cooperators on the grid. In the viscous case these cooperators had at least one neighbour with the same strategy (for $i > 1$), forming a connected patch of cooperators. For small numbers of cooperators this setting can be a good approximation of a patch of cooperators with one common ancestor, the initial single invader. In the well-mixed case, cooperators were placed randomly on the grid. After placing the $i$ cooperators, there could be two outcomes of the dynamics: either the number of cooperators increased by one after some time ($e_i \rightarrow ... \rightarrow e_{i+1}$), or decreased to zero ($e_i \rightarrow ... \rightarrow e_0$), where the expression ($\rightarrow ... \rightarrow$) can mean any number of steps (for example $e_i \rightarrow e_{i-1} \rightarrow e_{i-2} \rightarrow e_{i-1} \rightarrow e_i \rightarrow e_{i+1}$ for $i > 2$). Thus, the probability that the number of cooperators reaches the following
Figure 2.8: Invasion probability $\rho(x_0)$ with different benefit functions with well-mixed (on the left) and viscous (on the right) population structures. Triangle denotes the step, circle denotes the sigmoid and square denotes the linear functions. Curves are fitted for visualisation, the straight line represents the hypothetical neutral case. $c/b = 0.1$, $h_{cr} = 3$, for other parameters see the main text.

Elementary state, i.e. increases by one, is

$$p\{e_{i+1}\} = \frac{\sum (e_i \rightarrow ... \rightarrow e_{i+1})}{\sum e_i},$$

(2.3)

where $\sum e_i$ denotes all $(e_i \rightarrow ... \rightarrow e_{i+1})$ and $(e_i \rightarrow ... \rightarrow e_0)$ cases and $\sum (e_i \rightarrow ... \rightarrow e_{i+1})$ denotes those ones, where the number of cooperators increases by one.

Computing $p\{e_{i+1}\}, i = 1, 2, ..., 5$ by simulation in the viscous model, we found that the first elementary transitions (that is $e_1 \rightarrow e_2$, $e_2 \rightarrow e_3$,..) happened with a lower probability than in the neutral case (denoted by $p^\circ \{e_{i+1}\}$), but if the number of cooperators locally exceeded the threshold $h_{cr}$, the probabilities of the elementary transitions soon exceeded the neutral probabilities (Fig. 2.9). Furthermore, the difference between these probabilities does not necessarily increase with increasing number of cooperators. The reason of this phenomenon could be that the growing patch of cooperators starts to split to smaller patches, which behave just as the initial patch did. Thus, the key steps for a rare cooperator of invading a population of defectors are the transitions at the threshold $h_{cr}$. For comparison, at the well-mixed case, the probability of these transitions stays below the neutral (Fig. 2.9). We calculated analytically the probabilities of some transi-
tions in the case of the viscous model for $h_{cr} = 2$ (for details see Appendix B), and the calculated trends are consistent with the simulations: $p\{e_2\} < p^\circ\{e_2\}$, but $p\{e_3\} > p^\circ\{e_3\}$.

![Figure 2.9: Differences between the probabilities of elementary transitions of the VD and neutral games ($p\{e_{i+1}\} - p^\circ\{e_{i+1}\}$). The population is well-mixed on the right panel and viscous on the left. Square denotes the $h_{cr} = 2$, circle denotes the $h_{cr} = 3$ and triangle the $h_{cr} = 4$ cases, other parameters are the same as before.]

**Discussion**

Examining stochastic individual based models of the $N$-person Volunteer’s Dilemma we showed that while a rare mutant cooperator has very weak chance of spreading in a well-mixed system, its success in a viscous population is quite high. Thus, we can not only expect the stable coexistence of cooperators and defectors in the VD game, as has already been shown by models of well mixed populations (Archetti and Scheuring 2011), but the invasion of rare mutant cooperators is also highly supported in a viscous population that has stochastic selection dynamics. The invasion probability as a function of the initial concentration of cooperators follows a sigmoid curve in the well-mixed case and a concave saturating curve in the viscous case. The robustness of our result is supported by the fact that shapes of the curves remain characteristically the same even with varying parameters, such as the critical threshold, $h_{cr}$ or the cost/benefit ratio, $c/b$ (Fig. 2.4).
Similarly, cooperators spread successfully even if mixing is more intense (Fig. 2.7) and even if we use a smooth sigmoid benefit function instead of the step function (Fig. 2.8). These results complement somehow the studies using similar spatial systems (Chen et al. 2012b, Szolnoki and Perc 2010). These papers showed that the equilibrium fraction of cooperators reaches maximum at an intermediate threshold in the VD game (Szolnoki and Perc 2010) and at an intermediate inflection point and steepness with the sigmoid benefit function (Chen et al. 2012b). What they had not studied are the problem of invasion and the effect of mixing on the invasion.

In a viscous population, the invasion of cooperators is more probable as their number in a neighbourhood exceeds the threshold $h_{cr}$. These cooperators behave as nuclei of further invasion, and these nuclei grow further with a higher probability than collapse (Fig. 2.9). In this respect, the VD and the NPD are similar models, because cooperators in the NPD also need some kind of positive assortment to spread.

However, there is still an important difference between the two models: in the NPD, not only is the invasion of cooperation impossible without some kind of assortment, but also the coexistence of cooperators and defectors is (Hauert et al. 2006b), whereas in the VD, coexistence is possible without any kind of assortment. Our numerical results using different benefit functions support this statement. With sigmoid benefit functions, cooperation is able to spread in both the well-mixed and the viscous cases, just as in the model with the step function. When the benefit function is linear (NPD game), cooperation spreads successfully only in the viscous case. In the well-mixed case, there was no initial fraction of cooperators which would result in successful invasion inside the reasonable range ($x_0 \in [0; 0.3]$), which is in accordance with analytical and numerical results for two person PD games (Traulsen et al. 2006). Another important difference between the well-mixed and the viscous populations is that in the well-mixed case, the number of cooperators fluctuates strongly, so viscosity can help cooperators by lowering the magnitude of stochastic effects and consequently the probability of random drift driven extinction (Fig. 2.5). We cannot state, however, that viscosity generally helps the invasion of cooperation, since limited dispersal can increase not only the relatedness or spatial correlation among socially interacting individuals, but it may increase relatedness among potential competitors and intensity of competition, which might diminish cooperation.
(Van Baalen and Rand 1998, Hamilton 1971, Platt and Bever 2009, Taylor 1992, West et al. 2002, e.g.). Our lattice model is specific in this sense, since viscosity does not increase the intensity of competition, thus the positive effect of local relatedness is not overbalanced by more intense competition. Similarly, it is known that overlapping generations can have a positive or negative effect on the invasion of the altruist strategy (Lion et al. 2011, Platt and Bever 2009, Taylor and Irwin 2000). In our model, where generations overlap, vacant grid points emerge separately after the death phase if the death rate is not too large, thus, the intensity of competition is independent of the survival of individuals. Further, the reproductive success of individuals is also independent of the death rate, thus, as numerical simulations support, overlapping generations have no effect on the results.

Our numerical studies and calculations of the transition probabilities at the initial invasion support our intuition that the invasion of cooperators is successful in viscous populations if the number of cooperators exceeds the threshold in a local patch. This patch will continue to grow with a high probability. This is in accordance with the observation that the invasion probability function depends on the initial number of cooperators in the viscous case instead of the initial fraction. Initial aggregation naturally helps the spread of cooperators even in the NPD game, but the effect is much less pronounced.

Many conflicts in nature can be described by $N$-person non-linear games (Bednarz 1988, Packer et al. 1990, Creel and Creel 1995, Lee et al. 2008, Yip et al. 2008, Gore et al. 2009). Thus, reevaluating experimental and theoretical studies in the light of recent advances in the theoretical understanding of these games could move us closer to the understanding of social dilemmas in general. To give an example, our models have much in common with unicellular systems: the social dilemma has $N$ participants, the medium can be viscous, and the benefit function is probably sigmoid and highly non-linear. In light of the results given by our model, we can expect to observe the stable coexistence of cooperators and defectors in experimental populations with a similar structure. There are well-known empirical examples in the literature of public goods situations in unicellular populations (Gore et al. 2009, MacLean et al. 2010, e.g.), but these experimental works follow the tradition (along with many other) and apply the 2-person Snowdrift game instead of the more appropriate $N$-person non-linear game as a model framework to explain the experimental results. Both model systems predict the coexistence of cooperators and de-
fectors in concordance with the experimental results, but in an $N$-person non-linear game cooperators can invade only viscous populations easily, while in the 2-person Snowdrift game, cooperators easily invade well-mixed populations, too. Based on this difference, the applicability of these two model frameworks can be tested experimentally.
Chapter 3

Division of labour in a collective task

Behavioural differences among conspecifics of the same population have been reported from all kinds of living organisms, from bacteria to vertebrates, including humans (Dall et al. 2004, Dingemanse et al. 2004, Stamps et al. 2012, Grinsted et al. 2013, Cordero and 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 and Polz 2014), the adaptive value of individual variation in animal personality, especially in humans, is still the subject of debate (Penke et al. 2007, Montiglio et al. 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 and Taborsky 2010, Montiglio et al. 2013) and different forms of fluctuating or frequency dependent selection (Penke et al. 2007).

However, the possibility that selection for cooperation can also play a role in causing or maintaining this diversity, had 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 offspring (Clutton-Brock et al. 2004, Barta et al. 2014), group hunting (Stander 1992, Newton-Fisher 2014) or even trade in human context (Gowdy and 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.

In this chapter I introduce a strategic model that studies the effect of social division of labour (see Chapter 1) on heritable behavioural differences. As I have mentioned before, the first part of this chapter has already been published in Vásárhelyi et al. (2015). However, in the main text I only included results that are primarily my own work, details that are primarily my colleagues’ work, I included in Appendix C. The second, shorter part of the chapter discusses my newer, unpublished extensions of the model.

This study considers the necessary conditions and the potential evolutionary consequences of SDL. Using the cooperative task solving so typical in humans we try to estimate when and how SDL appeared in human societies. Furthermore, we aim to make a general argument about the possible role of SDL in creating and maintaining personality variation, an evolutionary process rarely considered by others.

Model outline

We consider 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 \( N_g \) individuals sampled randomly from the population. Individuals can choose between two kinds of subtasks, \( T_1 \) and \( T_2 \), 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 specialise more or less to one strategy, but then
they necessarily lose some of their flexibility: the more an individual specialises on one subtask, the less 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 $T_1$ or $T_2$ strategies, and the degree of specialisation to $T_1$ or $T_2$. For simplicity, we assume that the population is asexual, so progenies own the same trait values as their parent did. We characterise individuals with the strategy pair\(^1\) $x = \{p, c_1\}$, where $p$ is the probability of choosing strategy $T_1$ and $c_1$ is the cost assigned to it. Our main interest is whether DL emerges as a result of an evolutionary branching into two subpopulations specialised for the two subtasks.

Costs

As I have noted, each individual is characterised by the propensities of choosing each subtask and the costs assigned to performing them. Costs $c_1$ and $c_2$, associated with tasks $T_1$ and $T_2$, respectively, are constrained by the trade-off

$$c_1^\alpha + c_2^\alpha = k^\alpha$$  \hspace{1cm} (3.1)

(Fig. 3.1), where $k > 0$ incorporates 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, $T_1$ or $T_2$, with probability 1. The opposite is true when $\alpha < 1$. We will choose $c_1$ as the independent variable, therefore

$$c_2 = \sqrt[\alpha]{k^\alpha - c_1^\alpha}.$$  \hspace{1cm} (3.2)

Benefits

The total benefit of a group, $B(i)$, depends on the number of members ($i$) playing $T_1$, and consequently the number of members ($N_g - i$) playing $T_2$. $B(i)$ is a function typically with one maximum in $\{0, 1, \ldots N_g\}$. That is, there exists a distribution of $T_1$ and $T_2$ tasks

\(^1\)Note that this $x$ is different from the $x$ I have used in Chapter 2.
Figure 3.1: 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-$\overline{c}$ lines of given $p$ values (here $p = 0.4$) and different average costs. Higher the position of a line, the larger average cost, $\overline{c}$, it means. As we are looking for the lowest $\overline{c}$, we look for the lowest points where these iso-$\overline{c}$ lines touch the trade-off curves at least in one point. For $\alpha < 1$, we get the minimal $\overline{c}$ where the iso-$\overline{c}$ line is a tangent of the concave trade-off curve, the coordinates of the one common point define the optimal $c_1$ and $c_2$. The maximal extremum we get where the iso-$\overline{c}$ line is a tangent of the convex trade-off curve. For $\alpha \geq 1$, the minimum is one of the endpoints of the scale.

where solving the collective task brings the highest possible benefit. For simplicity we assume that $B(i)$ is symmetrical, i.e.

$$B(N_g - i) = B(i), \quad (3.3)$$

and has a single maximum at $N_g/2$.

We will study two types of models differing only in their way of benefit sharing. In model I), the group benefit, $B(i)$, is shared out equally among the group members, which could be a model of a cooperatively acquired common good from which nobody can be excluded. In model II), $B(i)$ is first split into two and then, within the subgroups playing
CHAPTER 3. DIVISION OF LABOUR IN A COLLECTIVE TASK

$T_1$ or $T_2$, 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 playing $T_1$ and $T_2$ are

\[
b_1(i) = \begin{cases} 
B(i)/N_g & \text{for model I} \\
\frac{1}{2}B(i)/i & \text{for model II} 
\end{cases} \tag{3.4}
\]

\[
b_2(i) = \begin{cases} 
B(i)/N_g & \text{for model I} \\
\frac{1}{2}B(i)/(N_g - i) & \text{for model II} 
\end{cases} \tag{3.5}
\]

**Predictions of the analytical models**

My colleagues had studied the above described model analytically (see Appendix C). Using adaptive dynamics (Geritz et al. 1997) they have made two main groups of numerically testable predictions, as follows.

**Possibilities of branching:** If $\alpha \geq 1$, the mixed strategy is never less costly for a given $p$, than specialisation. Therefore, the population will end up as being homogeneous with everybody specialised for one task and no one performing the other. Branching in this case will not be typical. In the special case of $\alpha = 1$ and $p = \frac{1}{2}$, selection is neutral. However, if $\alpha < 1$, for a given $p$, the mixed strategy is less costly than specialisation (that is, $p = 1$ or $p = 0$ for all individuals). Therefore the population will converge into $c_1 = \frac{1}{2}$. In the $x^* = \{\frac{1}{2}, \frac{k}{\sqrt{2}}\}$ point, branching can occur in certain ranges of the parameter $k$. Furthermore, branching will be much more frequent in model II), that is, where frequency dependence enhances diversity. For further details see Appendix C.

**The importance of the initial state:** We have pointed out that if $\alpha < 1$, $x^* = \{\frac{1}{2}, \frac{k}{\sqrt{2}}\}$ represents a singular point. However, in this case there must be at least two further singular points, too, at $x_u^* = \{p_u^*, c_u^*\}$, where $p_u^* > 1/2$ and $x_u^{**} = \{p_u^{**}, c_u^{**}\}$ where $p_u^{**} = 1 - p_u^* < 1/2$ (see Appendix C). This practically means that if we put the population in one of the end points, it will never converge into the central point where branching could occur (see also Fig. 3.2).

For further details on the mathematical analysis of the models, see Appendix C. In
CHAPTER 3. DIVISION OF LABOUR IN A COLLECTIVE TASK

Figure 3.2: Schematic figure capturing the singular points of the system. $x^* = 1/2$ and the $x^{(0)}, x^{(1)}$ end points are convergence stable points, and $x_u^*$ and $x_u^{**}$ are not. The latter two singular points separate the phase-space and make the dynamics converging either to $x^*$ or to one of the two end points. For further details see Appendix C.

the following I will concentrate on their numerical analysis. The building of numerical models is motivated by two main purposes. The first is to test the above predictions in a more realistic setting, that is, in a finite population with stochastic effects. The other one is to test the global behaviour of the model. We have seen that the mathematical analysis predicted the dynamics in the local neighbourhood of singular points, however, from an evolutionary point of view we have to see the gogal behaviour, too.

Individual based model outline

Our individual-based models were designed to be as similar to their analytical counterparts as possible. We present our methods and results using one arbitrary but illustrative parameter set concerning the size of the groups and the population, and the shape of 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 $x = \{p, c_1\}$. In each round, $N_g = 7$ individuals were sampled randomly from the well-mixed population of $N_P = 1000$ individuals to face the collective task. In each generation, 1000 such groups were selected, so that individual fitnesses were updated seven times on average 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 $B(i) = \frac{30}{\phi \sqrt{2\pi}} e^{-\frac{(i-\nu)^2}{(2\phi)^2}}$ with $\phi = 5$, $\nu = 3.5$ and $i$ denoting the number of group members.
playing one of the strategies. Varying $\phi$, 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 given 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 rounds of asexual 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 parent’s traits, but with probability 0.1, mutations occurred in both traits of the offspring. The mutant trait values followed a uniform distribution around the parents’ trait values, so they became $p' \in [p - \upsilon, p + \upsilon]$ and $c'_1 \in [\sqrt{c_1^{(\alpha)} - \upsilon}, \sqrt{c_1^{(\alpha)} + \upsilon}]$ with $\upsilon = 0.005$ (and with the obvious limitations that $0 \leq p' \leq 1$ and $0 \leq c'_1 \leq k$).

Results of the basic models

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 $p_0 = (1 - p_0) = 0.5$ and $c_0 = \sqrt{k^\alpha - c_0^{(\alpha)}} = \frac{k}{\sqrt{2}}$. This point, $x^* = \{\frac{1}{2}, \frac{k}{\sqrt{2}}\}$, is predicted by the analytical model to be the one, where branching could occur (for further details, see Appendix C). In each simulation we followed the population for a previously estimated time, enough for branching in most cases: $1.5 \times 10^5$ 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, although not with the original parameters, but evolving either to $x^{(0)} = \{0, k\}$ or $x^{(1)} = \{1, 0\}$

2) Branching occurred, half of the population became specialised in subtask $T_1$, and the other half in subtask $T_2$
3) The population was still distributed around the original states, probably meaning that there was not enough time for branching.

After mapping a certain part of the parameter space we concluded that the simulated results (Fig. 3.3) match the three most important predictions of the analysis. First, when branching occurred, it happened in or very close to the singular point (Fig. 3.4B). Second, the frequency of branching followed qualitatively the analytically calculated curves (see Appendix C, and the depiction in Fig. 3.3) in both models. Third, in model I) specialisation occurred in a much narrower area of the parameter space than in model II), as predicted by the analytical model. We also have a fourth observation, that is, in model II), above the predicted border line, the closer $\alpha$ to 1, i.e. the weaker the trade-off, the more branching occurs, probably as a result of the strong negative frequency dependent selection (see Fig. 3.3). Furthermore, while the invasion analysis is valid only in the close neighbourhood 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 $\alpha$ values as well, although with significantly smaller frequency than below the critical $\alpha$, because the $x^*$ singular point remains a branching point for every $\alpha$. 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 singular points $x^*_u$ or $x^*_u$ (see below).

Initial values and convergence

After testing the existence of the predicted singular point $x^* = \{1/2, k/\sqrt{2}\}$, we studied the effect of the initial population trait values, $\{p_0, c_0\}$, on branching in model I). As the analytical model predicted, the $x^{(0)}$, $x^{(1)}$ end points can be stable against the invasion of
mutants (see Fig. 3.4/A), that is, there is no branching when we use these initial trait values.

As we have pointed out before, the $x^* = 1/2$ singular point and the $x^{(0)}, x^{(1)}$ end points are convergence stable points, while $x_u^*$ and $x_{uu}^*$ 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 (see Fig. 3.2).

To test these predictions we ran simulations from $11 \times 11$ different starting points using the formerly tested parameter set ($k = 0.25, \alpha = 0.15$) that enabled branching. The results (Fig. 3.4) 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. Our numerically estimated results do not exactly match the predicted $x_u^* = \{p = 0.0154, c_1^o = 0.168\}$ and $x_{uu}^* = \{p = 0.984, c_1^o = 0.644\}$ points, nevertheless they are quite close (see Fig. 3.4/A). 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 (the behaviour of $x^*$) and during the simulations only extremely rarely did we observe them.

**Social control in the model**

We have already seen that in model I), which resembles the benefit sharing of egalitarian societies (Boehm 2001), branching, that is heritable specialisation, is less frequent than in model II), the negative frequency dependence of which represents “market” societies. However, I wanted to go further in modelling the egalitarian societies with introducing social control, and by decreasing the population size. Intuitively, social control means that there is a social pressure on every individual not to get specialised, or more practically, to carry out both subtasks with nonzero probability.

In the numerical simulations I incorporated social control in the following way: individuals still have the heritable parameters, $x = \{p, c_1\}$, however, they pick strategies according to a modified propensity, $0 \leq p^* \leq 1$. After the first occasion (when $p^* = p$), the propensity of choosing a subtask will be modified according to the previous round’s

Figure 3.4: The impact of the initial trait values on branching in model I). Plot A shows the initial parameter combinations with $k = 0.25$ and $\alpha = 0.15$. Darker points denote cases, 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.
choice. If an individual chose subtask $T_1$ in the previous round, then his propensity of choosing $T_1$ will decrease: $p^* = p - \chi$ and his propensity of choosing $B$ will increase: $(1 - p^*) = p + \chi$, where $0 \leq \chi \leq 0.5$ is the amount of social control.

For these simulations I have picked a parameter combination, where branching was guaranteed without social control, I used $k = 0.125$ and $\alpha = 0.2$. Otherwise, my parameters were similar to the ones used when investigating the branching behaviour of the model.

![Figure 3.5](image)

**Figure 3.5:** The average width of the branching fork as a function of an increasing social control. Darkening colours represent different times with yellow representing the initial conditions, orange representing generation 37 500, red representing generation 75 000 and black representing generation 150 000. $N_P = 1000$.

In Figure 3.5 I show the average width of the branching fork at different times with increasing social control. It can be seen that a small amount of control does not hinder branching, but definitely slows it down. However, above a certain amount of social control, the probability of branching decreases, which means that the population remains uniform with parameter values around the central singular point. The case is even more interesting when I decrease the population size. Figure 3.6 shows the same results with population sizes of 500, 250 and 150. It can be seen that in a smaller population social control hinders specialisation more seriously. The extreme case of $N_P = 150$ could seem arbitrary, but I have chosen it, because this is the natural social group size of modern humans (Dunbar 1993). It is quite surprising that in such a population size, even without social control,
the probability of branching is practically zero.

![Figure 3.6: The average width of the branching fork as a function of social control. Black dots represent the end of simulations, red ones the halftime. Plots differ in the population size. From left to right $N_P = 1000$, $N_P = 500$, $N_P = 250$ and $N_P = 150$.](image)

**Discussion**

We have highlighted how SDL 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 population size, the type of the benefit sharing, and the difficulty of invading the state where every individual is specialised 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 et al. 2012, Nakahashi and Feldman 2014), we believe that the SDL is an important new context.
We constructed a model framework for the emergence of consistent, heritable behavioural variation via SDL. We have shown 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 easily 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). Our numerical simulations have verified and extended the analytical results with more realistic conditions, such as a finite population and stochastic events. We have shown that there are parameter ranges, where branching is possible in the individual based model even if the mathematical model predicted it to be impossible.

We have seen that one of the biggest hindrances 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.

Finally, the last extension of the model has shown that branching can be further hindered by two additional factors, a social pressure against DL and a small population size. Interestingly, both of these circumstances could be typical in prehistoric times, as we will see in the next Chapter.

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 et al. 1990, Newton-Fisher 2014). However, we have little knowledge of a system where the members’
shares 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 et al. 2004, Kanngiesser and Warneken 2012). Obviously, this is not the only reason why we find far more examples of SDL 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.
Chapter 4

Food production and behavioural specialisation

Introduction

It is a common sense that the adoption of agricultural food production and sedentary life changed human societies immensely. This process, called the Neolithic transition, led to greater settlements, changing morals, social stratification and, later, to a number of other changes, like urbanisation and industrialisation (Diamond 1999, Boehm 2001, Gowdy and Krall 2013). It is less known that in different regions of the world the Neolithic transition happened independently at very different times (Diamond 1999). Agriculture first appeared at the Near-East, about 10–13,000 years ago, and soon it started its worldwide dispersal (Bar-Yosef 1998, Shennan et al. 2013, Boivin et al. 2016). Not much later it appeared in South-East Asia, then in the Americas, but there are societies that still did not (or did not fully) take up agriculture. These societies, which fall under the categories of small-scale societies, subsistence societies, hunter-gatherers, etc., are fairly good models of pre-agricultural (or pre-Neolithic) human life (Boehm 2000).

The most striking features of typical pre-agricultural societies include the absence of leadership, strong hierarchy, permanent homes and private property (Boehm 2001), all of which are present in post-agricultural societies. There are intriguing models treating the above aspects of the transition (Henrich and Boyd 2008, Bowles and Choi 2013),
CHAPTER 4. FOOD PRODUCTION AND BEHAVIOURAL SPECIALISATION

however, this study is concerned with two further outcomes of food production which affected human societies: food storage and social division of labour (SDL, see Chapter 1). In a certain sense social stratification in itself is a form of SDL, because a permanent DL naturally emerges between the strata. While DL by gender is present in all subsistence societies (Ember 1978), permanent DL between people of the same sex and age group is absent from mobile hunter-gatherer groups (Bird-David 2005, Gurven et al. 2009) and is only partly present in sedentary hunter societies (Boas 1895, Smith et al. 2003, Alvard and Gillespie 2004). Nakahashi and Feldman (2014) recently studied the evolution of DL in subsistence societies. They have shown that DL emerges if differences in resource acquiring and group size are large enough and there is food sharing within the group (see also in the Discussion). However, archaeological evidence suggests that SDL, either in the form of specialised individuals representing the first professions, or in the form of strong hierarchy, only appeared during the Neolithic (Wailes 1996, Finlayson and Warren 2012).

To continue the story here we are interested in how food storage, one of the earliest novelties of food producers, affects the emergence of SDL and phenotypic and genetic specialisation.

As we have seen in Chapter 1, DL is present at all levels of biological organisation, and is usually a sign of cooperation on a lower, and emergence of a higher level of organisation (Maynard Smith and Szathmáry 1995). Therefore, it is especially interesting that the Neolithic transition brought about a type of DL (that is, SDL) that was absent before. Our hypothesis is that the fundamentally different ecological environment of the Neolithic set the stage for widespread SDL that may have left a fingerprint on human genetics, too.

In this chapter I discuss a model that attempts to study the connections between the appearance of food production, human social division of labour and behavioural diversity. We have two main settings representing different social and ecological environments: the ones before, and after the adoption of agriculture. We study a social group of individuals who can choose and learn tasks, and who can imitate other, more successful individuals. The behaviour of an individual is affected by her genetics, her task choices, and whether she imitates others. This is in agreement with what we know about human personality: it is heritable to some extent with numerous genes, epistatic, and gene-environmental interactions involved, but otherwise is severely affected by the so-called “non-shared”
environment (Turkheimer 2000, Penke et al. 2007, Plomin 2011, Penke and Jokela 2016). Therefore, in the model we presume that behaviour has both a component of genetics and one of considerable plasticity. With the help of this model framework we attempt to answer the following three questions:

1) Can food storage and/or trade lead to SDL?
2) What are the necessary conditions and the limits for SDL to appear?
3) Does SDL induce only phenotypic, or genetic changes, too?

In the following we will first describe the model, then present and discuss our results and their consequences.

The model: general description

First we will give an overview of how the model works and then we will discuss each process in detail.

We consider a sexually reproducing, well-mixed group with group size $N_g$. In each time frame that we will call one round, individuals choose between $N_T \geq 2$ different tasks, and they spend their time with that particular task or activity. Being engaged in a task increases the payoff of the actor but tasks differ in the sense that they require and improve different skills. When someone spent a round with task $j$, this will not only gain her payoff, but also improve her effectiveness in solving task $j$. Individuals increase their expertise in one or more tasks during their lifetime. The more someone chooses a task, the more her expertise will grow in it, but the limit and speed of this increase depends on genetic factors (see Fig. 4.1). An individual’s genetics, $G_i$, consists of a quantitative genetic background (talent or affinity) for each task. We imagine this talent or affinity to be coded by a large number of genes, similarly to how personality or behavioural traits are coded by numerous genes (Penke and Jokela 2016).

At the end of each round payoffs are assigned to individuals according to their expertise in that round’s task. Occasionally, individuals have a choice to update their task choice strategy by imitating others. The purpose of such imitation is to copy successful strategies, but we assume that one is only willing to copy a strategy that is not too different from her own. The reason for this is that the complete change of the task choice strategy is rarely
more beneficial than the cost it bears, and we assume that individuals know this. At the end of a generation’s time (that includes a number of imitation phases, too), individuals are assigned into pairs and sexual reproduction takes place. Men and women are not distinguished in the model. The number of surviving offspring per couple is proportional to the average parental fitness. Offspring inherit genetics and task choice strategies from both parents.

In the following we discuss the above processes in detail.

**Task choice, expertise, genetics and payoff**

In each round, denoted by $t$, every member of the group chooses a task to spend her time with. Individual $i$ chooses task $j$ with probability $p_{ij}$, where $\sum_j p_{ij} = 1$. Offspring inherit $p_{ij}$ values from their parents, but these can change via imitation (see later). After individual $i$ solved task $j$, that is, at the end of the round, $i$’s payoff increases as follows:

$$b_{ij}(t + 1) = b_{ij}(t) + \beta E_{ij}(t),$$

where $b_{ij}(.)$ is $i$’s net benefit from task $j$ until round $(t+1)$ and $t$, $E_{ij}(t)$ denotes individual $i$’s expertise in task $j$ in round $t$, and $\beta$ is a constant that converts expertise into benefit. Naturally, in real life, the exact value of the payoff would scatter around an expected value, but for the sake of simplicity, we only use the expected value itself. Roundly payoffs per tasks are added together throughout an individual’s lifetime, and at certain times fitness is calculated from them (see later).

Expertise is a sigmoid function (Leibowitz et al. 2010) of both the time spent with a task and the genetic affinity ($G_{ij}$) assigned to it (see Fig. 4.1), according to the following equation:

$$E_{ij}(t) = \frac{G_{ij}}{1 + \exp\left(-\left(E_0 + \epsilon n_{ij} G_{ij} + \mu\right)\right)},$$

(4.1)

where $\epsilon$ is a constant, scaling the expertise function, $n_{ij}$ is the number of rounds individual $i$ has already spent with task $j$, $G_{ij}$ is individual $i$’s genetically inherited affinity for task $j$, $E_0$ is a constant basic expertise, and $\mu$ is a constant, positioning the inflection point of this sigmoid expertise function (see Fig. 4.1). So the genetic background plays an important
role in how one can improve in one task or another. Since the tasks can demand crucially different skills, as we will discuss later in detail, we assume a trade-off between talents for different tasks:

$$\sum_{j=1}^{N_T} G^\alpha_{ij} \leq N_T^\alpha,$$

(4.2)

where $N_T$ is the number of tasks, and $\alpha$ is a constant determining the strength of the trade-off. This trade-off function represents the idea that if someone is talented in a task, she will be relatively less talented in another task needing different skills, just as how people with different vocational interests tend to have unlike abilities (Randahl 1991).

The smaller $\alpha$, the stronger the trade-off. If $\alpha = 1$, the upper bound function of the possible $\sum_{j=1}^{N_T} G^\alpha_{ij}$ values in Equation 4.2 is linear, if $\alpha < 1$, it is convex and if $\alpha > 1$, it is concave. Thus we can study qualitatively different subcases of the model with different $\alpha$ values. If $\alpha < 1$, which we will call strong or stronger trade-off, then a greater talent for one task necessarily means a much smaller for another. If $\alpha > 1$, which we will call weak or weaker trade-off, then it is less costly for someone to have affinity for more than one task (Levins 1968).

**Fitness**

In the introduction we proposed this model to be able to show basic differences between two kinds of social and ecological settings, the ones before and after agriculture appeared. For this purpose we have examined the above model framework with two different fitness calculating procedures. The two different settings we will call the pre-Neolithic or subsistence, and the Neolithic or producer case. In the subsistence setting, because of the special ecological environment, individuals are better off when they have some expertise in all activities, while in the producer setting having a general knowledge or experience is no longer a necessity (for details see the respective sections below).

**The subsistence case**

In the subsistence case, we attempt to model societies that are similar to contemporary mobile hunter-gatherers. Most of these people, especially in warmer environments, move
Figure 4.1: Increasing levels of expertise for different genetic values. A higher genetic affinity grants faster learning and an altogether higher level of expertise (see also Eq. 4.1). Solid lines are calculated with the simulation’s initial genetic values, dashed ones with the corresponding theoretical maximums after complete specialisation. In other words, dashed lines represent higher, solid lines lower genetic affinities. The colours denote two main settings: blue lines represent stronger, red ones weaker genetic trade-offs (see also Eq. 4.2 and its explanation in the main text).

their camps several times a year, and do not have the opportunity for storing food or other resources on the long run (Binford 1980). As there is no food storage, but the limiting resources fluctuate strongly, each individual has to acquire different resources at different times. In other words, the ecological circumstances force people to gain expertise in all possible tasks or activities. Otherwise, there would surely be at least one phase in their life when the lack of knowledge and experience affected their fitness seriously. To ensure that individuals who lack at least one kind of experience will never be successful, we calculate the subsistence fitness as follows:
that is, we calculate the fitness as the product of the net benefit assigned to the different tasks (Sæther and Engen 2015). For a given \( \sum_{j=1}^{NT} b_{ij}(t) \) this function has a maximum where \( b_{i1} = b_{i2} = \ldots = b_{iNT} \). This will be attained most easily by a generalist individual, who has no special interest, nor outstanding expertise in any one task. We note here that as we are only interested in SDL, the model does not include gender-specific traits and corresponding genetic differences. Obviously, some sexual DL exists in all human societies, but it does not lie in the focus of the present study.

**The producer case**

In the producer case we attempt to model early agricultural societies. In contrast with the previous setting, here we assume that there is either large-scale food storage, frequent trade, or most likely, both (Bar-Yosef 2001). Producer individuals can be even more successful if they are occupied with one activity throughout their whole life, because they have a good chance of becoming experts of that task. We also assume that rare specialists are always better off than frequent ones, thus we incorporate the option of negative frequency dependence.

We calculate the specialist fitness as follows:

\[
W_i^{(p)} = \sum_{j=1}^{NT} \left( b_{ij}(t) \left( 1 - \delta - (1 - 2\delta)f_j(t) \right) \right),
\]

where \( f_j(t) \) is the frequency with which \( j \) was chosen till round \( t \), and \( \delta \in [0, \frac{1}{2}] \) is a constant, scaling the strength of the frequency dependence. The negative frequency dependence is maximal if \( \delta = 0 \), and there is no frequency dependence if \( \delta = \frac{1}{2} \).

**Imitation**

In an individual’s lifetime there are several phases of imitation. Then we draw two different individuals \( (l, m) \) randomly from the group, and imitation will occur with a probability given by \( \Phi(d_{lm}) \), where \( \Phi(x) \) is a Gaussian function with zero mean and variance \( \sigma_l^2 \).
\[ d_{lm} = \sqrt{\sum_{j=1}^{T}(p_{lj} - p_{mj})^2} \] is the Euclidean distance of the task choosing probabilities of individual \( l \) and \( m \). \( d_{lm} \) is smaller (and thus \( \Phi(d_{lm}) \), the imitation probability is larger for a given \( \sigma_I \)), if the task choice probabilities of \( l \) and \( m \) are similar, and \( d_{lm} \) is larger (whereas \( \Phi(d_{lm}) \) is smaller), if they are different. Thus, \( \sigma_I \) determines the strength of the assortativity: the smaller \( \sigma_I \), the stronger the assortativity. Assortativity means that individuals are more likely to copy others who are similar to themselves regarding their task choice strategies. Again, similarity here does not refer to gender-specific traits, but other behavioural traits. We repeat this algorithm as long as we do not find enough \((l, m)\) pairs between whom imitation occurs.

When a pair, \( l \) and \( m \) was chosen, \( m \) will copy the task choosing preferences of \( l \) with probability

\[
\left(1 + \exp\left(-\gamma \left(W^{(i)} - W^{(m)}\right)\right)\right)^{-1},
\]

where \( W^{(i)} \) is the fitness of individual \( j \) either in the subsistence or in the producer scenario, and \( \gamma \) is a constant that determines the strength of the selection. If \( W^{(i)} > W^{(m)} \), it is more probable that \( m \) will copy \( l \), and vice versa (Szabó and Fáth 2007a).

If \( m \) imitates \( l \), \( m \)'s updated probabilities, \( p'_{mj} \), of choosing task \( j \) will be drawn from a normal distribution \( \mathcal{N}(p_{lj}E_{mj}(t), \sigma^2_p) \), where \( \sigma_p \) scales the precision of the copying. As the new probability scatter around the product of \( l \)'s probability and \( m \)'s expertise, \( m \) will rarely switch to a strong preference for something she has no experience in (that is if \( E_{mj}(t) \approx 0 \)). Note that if assortativity is strong, individuals are prevented in two different ways from copying strategies that are far from their own. First, \( m \) will rarely switch to task \( j \) if she has no prior experience in it (that is, if \( E_{mj} \) is small). Thus \( m \) can avoid decisions, where the benefits are far below the costs. Second, when \( m \) copies \( l \), \( m \) will get a mixture of the model’s strategy and her own interest. An individual is involved on average in five imitation events during her lifetime, so potentially she can end up with a quite different task choice strategy from the one she had been inherited.

**Reproduction**

At the end of a generation’s time, reproduction takes place. Individuals reproduce sexually which means the recombination of traits. Similarly to imitation, reproduction can also
be assortative with regard to task choice strategies. As we have mentioned before, the assortativity of imitation is based on the assumption that it would be too costly to totally change one’s task choice strategy if she already has expertise in something. The rationality of assortative reproduction is less obvious. Yet, there is evidence that people tend to prefer and choose similar pairs regarding several characteristics, and that some of these similarities are associated with a higher reproductive success (Huber and Fieder 2011, Krzyżanowska and Mascie-Taylor 2014, Conley et al. 2016).

At the beginning of reproduction, individuals \(l\) and \(m\) are selected randomly and paired with a probability given by a Gaussian curve, with zero mean and \(\sigma_R^2\) variance, at the task choice strategy differences \(d_{lm}\) (see above). The larger \(\sigma_R\), the weaker the assortativity, thus a very large \(\sigma_R\) gives back random mating. The number of offspring a pair will produce is proportional to the average relative parental fitness. An offspring’s genetic values, \(G_j'\), are drawn from \(\mathcal{N}\left(\frac{1}{2}(G_j^\times + G_j^\circ), \sigma_G^2\right)\), where \(G_j^\times\) and \(G_j^\circ\) are the parental values and \(\sigma_G^2\) determines the variance of inheritance. We assume here that genetic affinities are determined by the small effects of numerous loci and that these loci act dominantly in an additive way. The Gaussian distribution of genetic values is the approximation of the Mendelian segregation of these additive alleles. While this assumption is under debate for years (Hansen 2013, Nelson et al. 2013), it has recently been shown by meta-analyses of twin studies (Hill et al. 2008, Polderman et al. 2015) and by further theoretical models that additive genetic variance generally dominates complex traits in natural outbred populations (Hill et al. 2008, Mäki-Tanila and Hill 2014). If, by chance, \(\sum_{j=1}^{N_f} G_j'^\alpha > N_f^\alpha\) (see Eq. 4.2), we choose a random \(G_j'\), decrease it by 5%, and repeat this as long as necessary. The offspring’s task choice strategy is simply determined by the average parental probabilities, \(p_j' = \frac{1}{2}(p_j^\times + p_j^\circ)\). Our individual based model is related to the adaptive dynamical description of frequency dependent selection (Geritz et al. 1998) but it is more general than that by enabling mutants being at any distance to the resident type, using assortative mating (Dieckmann and Doebeli 1998) and a small group size where drift can have a significant role (Waxman and Gavrilets 2005).

Offspring produced by the pairs get into the offspring pool, from where we randomly choose \(N_g\) individuals for the next generation. Sometimes there was not enough offspring to fill up a group of \(N_g\), mostly because there were one or few individuals with such a high
relative fitness that others hardly produced any offspring. In these cases we repeated the
t reproduction phase as many times as necessary.

Simulation settings and questions

During the basic simulations we use a social group of \( N_g = 100 \) individuals. We initialise
the group in a way that every member has the same genetic affinities, \( \forall i, j : G_{ij} = \sqrt{\frac{N^2}{N_T}}, \)
that is, everybody has equal talent for all activities. Importantly, this does not mean that
everybody has the exact same genetics, only that they have identical quantitative traits.
Furthermore, everybody has the same task choice strategy, \( \forall i, j : p_{ij} = \frac{1}{N_T} \). Then we run
the simulations for \( 10^3 \) generations. Each generation lives for \( 10^3 \) rounds, during which 10
imitation phases take place with \( N/4 \) imitation events each. At the end of a generation’s
time we record the \( G_{ij}, p_{ij}, \) and the \( W_i^{(l)} \) values in the group.

Results

Results with two tasks

We have picked a basic parameter set with which we can demonstrate the main results and
the differences between the subsistence and producer cases. These parameters we tried
to set as realistic as possible, but also to give straightforward and illustrative results. In
the following we will only note parameter values that are different from the basic values
listed in Table 4.1. Our primary question is whether the optimal behaviour in the two
settings are different. We also ask whether the behaviour has an effect on the genetics,
and whether the strength of the genetic trade-off alters these results.

It can be seen in Figure 4.2/I and 4.2/II that with this basic parameter set our two
settings show striking differences, and the results are not affected qualitatively by the
strength of the trade-off within this range.

In the subsistence case we can see that group members are generalists both on the
phenotypic and the genetic levels (see I/A and II/A in Fig. 4.2). This is so with both
weaker and stronger trade-offs, though there are more specialists in the group when the
trade-off is stronger. As in the subsistence case specialists have a very low fitness compared
Figure 4.2: Summary of the basic results in both settings with stronger (I) and weaker (II) trade-offs. The figure shows the task choice probabilities (top row) and the distribution of the genetics (bottom row) for one task throughout all generations. The top right plots show the fitness of individuals as a function of their average life-long task choice probabilities. The bottom right plots show these average probabilities as function of the genetic affinity assigned to the same task. The greenish colours (A columns) represent the subsistence, the reddish (B columns) the producer case. For the parameters see Table 4.1.
CHAPTER 4. FOOD PRODUCTION AND BEHAVIOURAL SPECIALISATION

Table 4.1: Parameters of the simulations. The Value column contains the basic parameter set. The Range column shows the parameter ranges that we studied throughout the simulations. SD stands for standard deviation.

*In all cases of $\epsilon > 0.04$ we used $\gamma = 5$ to avoid a numerical overflow error.

to generalists (see the top right plots in Fig. 4.2/I and 4.2/II), the appearance of specialists is not a product of selection, but more an effect of the trade-off and the process of imitation. When the trade-off is stronger, the sum of the genetic affinities of a non-specialist individual is lower (see Eq. 4.2), therefore it is more difficult to become expert in both tasks. Meanwhile, imitation is strongly assortative, therefore a couple of specialising individuals are more likely to imitate each other and get even more specialised. Note that the individual who is the model of imitation does not need an absolute high fitness, it is enough if she is fitter than the imitating one. Furthermore, though most individuals are generalists, there still is a positive correlation between the genetic affinity and the probability of performing the same task (see the bottom right plots in Fig. 4.2/I and 4.2/II).
In the producer case we see specialisation on both levels (see I/B and II/B in Fig. 4.2). However, while on the phenotypic level everyone is completely specialised to one task, on the genetic this is not so. It seems that there is a constraint on specialisation in this latter case: in spite of the branching on the phenotypic level, complete specialisation on the genetic level is absent with both weaker and stronger trade-offs. Nonetheless, there is a strong correlation between the genetic affinity and the probability of performing the same task (see the bottom right plots in Fig. 4.2/I and 4.2/II). Note also that the two groups of specialists have different mean fitness values (see the top right plots of Fig. 4.2/I and 4.2/II). This temporary difference is caused by stochastic fluctuations and frequency dependence, and disappears in the long run.

It is important to note here that all of the plots in Figure 4.2 are generated using the values assigned to only one task. As the same plots look quite similar for the other task, we decided not to include them.

As the group size of $N_g = 100$ is not realistic as an effective population size (see later), we have tested the basic model for a much larger group size, $N_g = 1000$, too. These latter results look quite similar to the basic ones, only enlarging the group lessens the noise in the distribution of traits (see Fig. S3). Therefore, in the following we will keep using the $N_g = 100$ case for most model checks, and assume that this focal group is part of a larger population.

The strength of assortativity

During our basic simulations both the imitation and the reproduction were strongly assortative. Theoretically, strong selection could have the same effect as strong assortativity, because individuals accidentally imitating peers with more similar task choice strategies are more successful. In practice, however, if imitation is random, it has a strong mixing effect on the task choice strategies and thus weakens the effect of selection significantly. Therefore, we decided to put strong assortativity into the basic model. Nevertheless, we need to ask the question, what happens if assortativity is weak or even absent. During the following we will present the results of the producer case, because the lack of assortativity does not qualitatively affect the subsistence case.

We have tested our model with weaker ($0.03 << \sigma \leq 1$) and with no assortativities
and the results are as follows. If strong assortativity is present during only one of the life events (imitation or reproduction), while during the other assortativity is weak or even absent, then phenotypic specialisation is still typical, but genetic branching does not occur. However, if assortativity is absent both during imitation and reproduction, the results become very different. With stronger genetic trade-off group members soon end up as specialists of only one task with very low fitness, that is, drift drives the group to a suboptimal stable state. However, if the trade-off is weaker, the genetic backgrounds in the group remain generalist-like, and thus there is no drift on the phenotypic level, either. Weaker genetic trade-off means exactly that it is possible without a large cost to have genetic affinity for more than one task, hence the absence of drift in this case. We note here that the correlation, mentioned in the previous section, between the genetic value and the probability of choosing a task is still present here. Therefore, even if there are no genetically distinct subgroups, the genetics–phenotype relation still defines distinct behavioural types.

In summary, if either imitation or reproduction is strongly assortative, phenotypic specialisation likely occurs, but for genetic branching, we need strong assortativity at both life events (see Table 4.2).

<table>
<thead>
<tr>
<th></th>
<th>(\alpha = 0.75) stronger trade-off</th>
<th>(\alpha = 1.25) weaker trade-off</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>phenotypic</td>
<td>genetic</td>
</tr>
<tr>
<td>random mating</td>
<td>Sp.</td>
<td>(\varnothing)</td>
</tr>
<tr>
<td>random imitation</td>
<td>Sp.</td>
<td>(\varnothing)</td>
</tr>
<tr>
<td>random mating &amp; imitation</td>
<td>drift</td>
<td>drift</td>
</tr>
</tbody>
</table>

Table 4.2: Summary of the effects of strong assortativity in the producer case. Sp. means specialisation, \(\varnothing\) means “no specialisation”, and drift denotes the case where all group members become specialised in the same task.
CHAPTER 4. FOOD PRODUCTION AND BEHAVIOURAL SPECIALISATION

The speed of learning

So far we assumed a relatively fast learning. We suggested that a genetically generalist individual (with a stronger trade-off) can achieve a nearly maximal expertise within a fifth of her lifetime, that is, within 200 rounds, if she spends every day with the same task (see the solid blue line in Fig. 4.1). If she is genetically specialised to that task, too, she can master the task even sooner (see the dashed blue line in Fig. 4.1). In Neolithic human terms, it would mean that any of the two activities can be mastered within approximately 5 years of adult life (Galor and Moav 2007), which does not seem to be a far-fetched assumption for moderately difficult tasks. However, as some tasks can be far more difficult to learn, we also studied our model with the assumption that learning a task is slower.

In concordance with the intuition, when it is slower to master a task, it is more difficult to be a generalist. If we decrease the learning speed, a couple of specialists appear at first in the subsistent group. If we decrease it even further, the generalists disappear on the phenotypic level, but on the genetic level we do not see any branching. In this case all members’ fitness is very low in the subsistence group. In the specialist case we see nothing unusual: everybody is specialised to one task or another, and genetic branching occurs, too.

To summarise, we see a general pattern that if tasks become difficult to learn, specialisation is inevitable on the phenotypic level. This corresponds well to the everyday experience: if tasks are difficult or skill intensive, division of labour gives a higher return.

The role of frequency dependence

So far we assumed that the negative frequency dependence of the producer fitness was very strong ($\delta = 0.01$). Thus, we tested the model for weaker and for no frequency dependence with the following results. If frequency dependence was absent, we saw drift on both the phenotypic and genetic levels, similarly to the case where assortativity was absent both during imitation and reproduction. This is not surprising, and has some interesting relevance.

If there is no frequency dependence, it means that a group as a whole can be successful
when everybody is specialised to the same task. This is possible in real life too, if a group
has frequent opportunities for trade, and has something to sell for a good price. Then
a specialised group can easily maintain itself from its only activity (or product). Such
an example could be the Polynesian settlement on the Pitcairn and Henderson islands,
where life is only sustainable if trade supplies many necessary resources (Diamond 2005).

If frequency dependence was weaker, but not absent, for a surprisingly wide range ($\delta \leq 0.35$) it did not affect the process of genetic branching and the corresponding phenotypic
specialisation. Thus, we can say that the producer group needs some, but not necessarily
strong negative frequency dependence to have two distinctly specialised subgroups.

**The mechanisms generating the variance**

In this model we have three types of variance generating events: the vertical transmission
(inheritance) of $G_{ij}$, the vertical transmission (inheritance) of $p_{ij}$, and the several hori-
zontal transmissions of $p_{ij}$ (imitation of peers). In the basic parameter set the genetic
transmission is much less accurate than any of the task choice strategy transmissions.
However, as we have only one vertical transmission and several horizontal transmission
events, individuals have a large potential to end up with a different phenotype from what
they had inherited. Whether this potential is realised with the basic parameters depends
both on the setting and the strength of the trade-off. Nevertheless, according to our mea-
sure, the standard deviation of the difference between the inherited and final phenotypes
can be even a magnitude higher than the standard deviation of a single transmission.

We have conducted tests to find out how much these parameters affect our general
results. We have found that at least one of the three variance generating events has to be
inaccurate enough to get back the typical results on the phenotypic level. If the genetic
transmission is highly accurate (e.g. $\sigma_G = 0.01$), there will be no genetic branching and
as $\sigma_G$ gets larger, the branching gradually gets more and more pronounced. Importantly,
if all transmission events are inaccurate, then the distributions are more noisy, but qual-
itatively the same, that is, decreasing the accuracy of imitation does not affect our results
qualitatively. Therefore, using smaller $\sigma_p$ values does not change the results, only makes
them less noisy.

Finally, we have tested our model with an exponentially decreasing $\sigma_G$ based on the
assumption that the assortative reproduction decreases the additive genetic variance on
the long run until it reaches an equilibrium. In this test we have found that it is only the
endpoint of the decrease that affects the results in the above explained manner, but the
decrease in itself has no effect.

To summarise, for genetic branching to occur, we need at most a moderately accurate
genetic transmission, but with such a $\sigma_G$, the accuracy of the imitation does not affect
the results qualitatively.

**Results with more than two tasks**

When we studied our model with more than two tasks, we consequently faced the problem
of scaling. Consider for example a generalist group, where people master all the existing
tasks throughout their lifetime. Obviously, if the number of tasks is four instead of two,
everyone needs twice as much time to master all the tasks, if all other parameters are
unchanged. But if it takes so much time to learn everything, an alternative strategy may
gain higher return than the pure generalist. Also, the probability of randomly choosing
all tasks nearly equally or equally often in the first couple of rounds is decreasing fast
with the number of tasks. For example, for 12 consecutive rounds of task choice, the
probability of even choice for two, three and four tasks, according to the multinomial
distribution, are $0.2256$, $0.0652$ and $0.0220$, respectively. Then, if one task gets chosen
more frequently, by chance, this will trigger a positive feedback loop, where the preference
for the practised task will push the individual towards specialisation, even if remaining a
generalist would have been more beneficial. Hence, simply increasing the learning speed
does not solve the problem. Therefore, in the next section we will not attempt to compare
results with previous ones explicitly, but only to give an overview of the trends we saw.

First we introduced a third task into the model and did not change any other param-
eter. When the trade-off was stronger, in the producer case one task disappeared, and
then individuals specialised to one of the two remaining tasks both phenotypically and
genetically, similarly to the $N_T = 2$ case (see Fig. 4.3/A). In the subsistence case we also
saw specialisation, but only on the phenotypic level, and consequently in this group ev-
everyone had very low fitness. When the trade-off was weaker, the results were similar with
the difference that in approximately half of the cases no groups of specialist disappeared in the producer case (see Fig. 4.3/B). Also, in the subsistence case there were a couple of individuals who were able to stay generalists. The latter had very high fitness compared to others, so the next generation practically contained only their offspring.

Because of the previously mentioned demographic stochasticity, it is worth studying the problem of specialisation with larger group sizes, or in a meta-population, too, if there are more than two tasks. Therefore, we have created an extension for the model with a meta-population structure of 10 groups containing 100 individuals each. After a generation’s time, and just before reproduction, we have randomly chosen a fixed proportion of individuals (1–10%) from every group for migration. Then we assigned random pairs of groups exchanging their migrants. Thus we have created a constant gene flow among the different groups, while all individuals spent most of their life in only one group.

In agreement with our expectation, both the meta-population structure and the larger
group size \((N_g = 500 \text{ and } N_g = 1000)\) helped genetic specialisation, while phenotypic specialisation was also complete. In the case of the meta-population structure, even the smallest migration of individuals prevented the specialist subgroups from disappearing. However, there were still no generalists in the subsistence case. In the case of the larger groups, we have noted similar trends. As we increased the group size, the quantitative genetics of the specialised groups became more distinct with both trade-offs in the producer case (Fig. 4.3/C), and the three specialist groups coexisted stably throughout the simulations. The subsistence case was somewhat different though. When the trade-off was stronger, individuals specialised, but when it was weaker, there were both a number of specialists and many generalists with high fitness in the group (see Fig. 4.3/D). Note that in the two ecological environments different strategies (and genetics) are associated with higher fitness, and this is well captured in Fig. 4.3. In the subsistence case points representing high-fitness individuals (the generalists) scatter around the middle. In the producer case, points representing high-fitness individuals (the specialists) are the ones closest to the vertices.

Finally, we studied the subsistence case with increased learning speed. In agreement with the intuition, if we increased the learning speed significantly, the subsistence group did not get specialised. With intermediate learning speed we saw that the strength of the trade-off determined whether a generalist strategy was viable or not: weaker trade-off enabled most individuals to stay generalists, but stronger trade-off made them more specialised (see Fig. S5 in Appendix D). Consequently, the fitness of group members was much higher in case of the weaker trade-off \((\alpha = 1.25)\).

To summarise, when there are three tasks that are moderately slow to learn, a group of 100 is not big enough for complete and stable specialisation. Also, the cost of learning three tasks instead of two drives the subsistence group into specialisation and to a very low fitness state, unless the speed of learning increases, too.

After studying the model with three tasks, we introduced a fourth one. In case of four tasks an even larger group size is needed for coexistence of four subgroups of different specialists. We have studied the \(N_g = 500, N_g = 1000\) and \(N_g = 2500\) cases and the following results qualitatively hold for \(N_g \geq 1000\). As it is shown in Figure 4.4, phenotypic specialisation is almost complete and there is a clear direction towards genetic specialisation.
with weaker trade-off, even if the specialist groups are less distinct than in the previous cases. However, when the trade-off is stronger ($\alpha = 0.75$), the specialisation is not stable, one task typically disappears, similarly to the $N_T = 3$ case (see Fig. S6 in Appendix D).

With the subsistence setting we have a similar situation as before. When the trade-off is stronger, it is far too difficult to remain generalist, therefore only a very small number of individuals end up with high fitness (see Fig. S6 in Appendix D). However, when the trade-off is weaker, there are a great amount of generalists in the group (see Fig. 4.4).

In summary, as we increase the number of tasks, it is more and more difficult to either get completely specialised in the producer, or remain truly generalist in the subsistence case. The more tasks are present, the larger group is needed for stable specialisation. Note, however, that the two settings are quite different from each other qualitatively. Weaker trade-off helps both stable specialisation in the producer case and the maintenance of a generalist group in the subsistence case.

When there are several tasks, it seems to be a particularly strong assumption that all tasks are completely independent regarding their genetic background, therefore we designed a further variation of the model with overlapping genetics.

**Overlapping genetic backgrounds**

In this version of the model we assumed that while the four tasks need four different kinds of expertise, they have only two underlying quantitative genetic backgrounds. That is, both $E_{i1}$ and $E_{i2}$ are function of $G_{i1}$, and both $E_{i3}$ and $E_{i4}$ are function of $G_{i2}$. This way, if someone is born with specialist genetics, she can still choose between two tasks to devote her life to. We studied this model version with $N_g = 2500$ and different speeds of learning. With a lower speed ($\epsilon = 0.2$ that is still higher than the original) we found that in the producer case there was complete phenotypic and clear genetic specialisation (see Fig. S7 and S8 in Appendix D). This means that on the genetic level the group divided into two subgroups, and on the phenotypic level, each of these subgroups divided into two groups of specialists. However, if we increased the learning speed even further, to $\epsilon = 0.8$, where generalists could appear again in the subsistence setting, the genetic specialisation in the producer case disappeared.
Figure 4.4: Phenotypic (top) and genetic (bottom) profiles in the $N_T = 4$ case with weaker trade-off ($\alpha = 1.25$). The figure captures tetrahedron nets, where each triangle is a simplex calculated from the three noted tasks, similarly to the simplexes in Fig. 4.3. Thus, the faces of the tetrahedrons show projections of the 3D mass of points. Each point corresponds to an individual, and plots contain all group members. Reddish colours represent the producer, greenish the subsistence case. For further parameters, see Table 4.1.
Moreover, we ran simulations, where we introduced two further tasks with a shared genetic background, but with a group size of $N_g = 5000$ we still not had six stably specialised subgroups, because some tasks disappeared. Unfortunately, we could not study larger groups and numbers of tasks because of computational time limits.

To summarise, if we assume that tasks have overlapping genetic backgrounds, with 4 tasks both phenotypic and genetic specialisation can occur at a group size of 2500 individuals, but further tasks probably require even larger group size for specialisation.

**Discussion**

Given the existing differences between the ecological environments of subsistence and producer groups, it is not surprising that SDL is rarely present in the former while widespread in the latter. However, a more detailed study of such a consequential problem is still necessary if one wishes to understand the radical change that followed the Neolithic transition. Our model framework incorporates two significant aspects of proficiency: the genetic background and the phenotypic plasticity of the propensity and efficiency of task solving, that is, individual learning and imitation. Furthermore, the model is suitable to study the effect of group size and the strength of genetic trade-offs. With this comprehensive model we have a chance to understand how the varying characteristics of individuals and groups affected the appearance of SDL and specialisation during the Neolithic transition.

In the following we will discuss the results and implications of the subsistence and producer settings separately.

**The subsistence case**

Our subsistence case intents to model a mobile hunter-gatherer group. Individuals have to gain experience in all tasks, just as hunter-gatherers have to participate in all activities of the group (Marlowe 2010, Boehm 2001). We have not included all possible fitness affecting factors into the model though. For example, while food sharing would have decreased payoff differences (Kaplan et al. 1984, Marlowe 2010) thus helping the less successful, reputation would have increased fitness differences (Kaplan and Hill 1985) thus counteracting the effect of food sharing. We decided that instead of further increasing
the complexity of the model we assume to capture roughly all of these effects with the fitness function. Furthermore, we did not incorporate external determinants of behaviour into the model. There are numerous narratives about how hunter-gatherers shape each other’s behaviour by means of teasing, mocking, or in extreme cases, ostracising (Boehm 2001). This culturally transmitted system of social pressure can help younger individuals to find the “optimal” behaviour in a given social and ecological environment. (I note here that this is exactly what I incorporated in the previous study as “social control”.) In the context of this model, such a mechanism could help individuals in the subsistence group to avoid the positive feedback loop of specialisation.

In our model there are two conditions that are necessary for the subsistence group to remain generalist: learning a task have to be relatively fast and there have to be at least one moderately large variance generating mechanism. If there are more than two tasks and learning a task is not easy enough, the members of the subsistence group specialise to certain tasks and have very low fitness. The reason for this is combinatorial: with linearly increasing number of tasks the probability of choosing all of them equally often decreases not linearly, but faster. Therefore, chance leads generalist individuals to preferences, and initial preferences easily drive these individuals into specialisation. If the task learning speed would increase with the number of tasks, the subsistence group could remain generalist. However, it is not obvious, whether and how the task learning speed could be a function of the number of tasks one learns. One argument could be that each learning process gives something to another one, therefore the more tasks one learns, the better she gets in all of them. For example, some studies find that children with music training outperform their peers in one or more other disciplines (dos Santos-Luiz et al. 2016). However, there can be another argument against this one: if tasks are significantly different, there could be a task switching cost (Goldsby et al. 2012) or simply different tasks could need mutually exclusive skills. Therefore, we can conclude that according to our model, learning a number of tasks and remaining generalist is only possible, if learning is fast or if the tasks need overlapping skills. For example, mobile hunter-gatherers are truly generalist, they participate in all tasks, such as hunting, gathering, weapon making or collective decision making (Marlowe 2010, Boehm 2001), but these tasks may share the need for several skills, such as cooperativeness or patience. But in some other small-scale
societies weak or temporary SDL appears, when new skills are needed. One example could be the specialisation for weapons or foraging techniques when the opportunities are wider (Dwyer and Minnegal 1991), another example the temporary DL during the season of agricultural work (Stieglitz et al. 2013).

A further aspect of the above is that it may add to the explanation of the so-called Sapient Paradox, that is, why the behaviourally (and probably genetically) modern humans had no technological revolution between the Upper Palaeolithic and the Neolithic (Renfrew 2008). The *Homo sapiens* went through several revolutionary changes during its evolution, however, between its cognitive explosion about 40 000 years ago and the appearance of agriculture, surprisingly few novelties appeared (Renfrew 2008). According to our results, until the ecological environment did not allow large groups, specialisation and division of labour, behavioural diversity must have been lower, and innovation scarcer. Interestingly, the oldest European remains of representational art (e.g. cave Altamira of Lascaux) are relatively close to the sea that serves as a much more reliable source of food than terrestrial habitats. Indeed, contemporary marine hunter societies have both *some* SDL and specialisation (Smith et al. 2003, Alvard and Gillespie 2004) suggesting that marine hunting is somewhere between mobile hunter-gatherer and producer ecologies.

**The producer case**

Our basic parameter set enables the producer group to part into specialised subgroups which engage in SDL. However, not every parameter set is ideal for such specialisation, especially not for genetic branching. One absolute need for specialisation on both levels is a relatively strong assortativity during imitation and reproduction (see Table 4.2). If only one of these events is strongly assortative, genetic branching is absent, but phenotypic specialisation is still present. Another necessity for genetic branching is a big enough group or population size. This is not surprising: if the group is small and not connected to other groups, each specialised subgroup is in danger of disappearing due to stochastic events. In the case of three tasks and a small group, for example, one task disappears in the producer case, but not in the subsistence case, while both group’s average fitness is relatively low. Interestingly thus, in a small group the only chance for solving all
difficult tasks is to value generalist tendencies even if eventually everybody ends up being a specialist. (If tasks are faster to learn, the case is similar, only the subsistence group ends up consisting of generalists and having a higher average fitness.) A final necessity is a large enough variance when inheriting genetics. This is in agreement with the previous point, too, as the equilibrium variance a population can reach depends on the size of the population itself. To summarise, as the number of tasks grows, an ever larger group or a well-connected meta-population is needed for maintaining specialist groups, and the necessary size of the group increases very fast with the number of tasks. This pattern corresponds to the general finding that in several animal societies DL is more pronounced in larger groups (Cartwright 2003, Thomas and Elgar 2003, Jeanson et al. 2007). It is also in agreement with the somewhat connected observation that larger group or population sizes contribute to larger cultural complexity (Henrich 2004, Derex et al. 2013).

Naturally, if tasks have overlapping genetics, the problem partly gets back to the ones with only two or three tasks. In these cases a medium size group is also big enough for several groups of genetic and phenotypic specialist to coexist. Again, this assumption that different task types share the necessary genetic background is not far-fetched. Think about it the other way round: those who share a genetic background will sometimes choose the same professions, and often similar ones. Indeed, several twin studies have shown that the variation in vocational interest is significantly influenced both by genetic and environmental factors (Schermmer and Vernon 2008), similarly to personality.

This model setting with the overlapping genetics seems more realistic than the ones without it, nevertheless, we think that it captures an extreme situation, too. We intend to investigate a model where the amount of genetic overlap behind tasks is also a variable parameter.

The limitations of our study

The model discussed in this chapter examines a problem rarely addressed by others, but as such pathfinder works, it has several limitations.

Some limitations arise from decisions that render the model more tractable. These include non-overlapping generations, or a stable group size, obviously neither of which
applies to humans. However, these restrictions or simplifications probably do not alter the results fundamentally.

Another simplification is that we only focus on one social group of $N_g = 100$ individuals, which is the cultural unit, and assume it to be part of a larger reproductive pool, a meta-population. Thus we assume that the genetic variance is larger than it would be if the effective population size were $N_g = 100$. We have shown that this assumption does not alter most results qualitatively, however, a reasonable further step could be to generalise our model with an extended meta-population structure, similar to the one we have discussed before. In such a model, it would also be possible to evaluate the role of competition among groups, that is, the role of multilevel selection in the evolution of SDL.

Finally, as we have already mentioned before, this simple model did not contain some important elements of the typical subsistence life, such as sexual DL, food sharing, reputation based mate choice or social pressure. We believe though that our fitness calculating procedure roughly captured the main effects of all the above.

**The Neolithic transition and human behaviour**

During the Near-Eastern Epipalaeolithic and early Neolithic, climate fluctuations became larger, human density grew, and the populations of the megafauna became scarcer (Diamond 1999, Bar-Yosef 2001, Simmons 2012). At the same time, human groups gradually changed their subsistence from mobile hunting and gathering towards sedentary food production (Maher et al. 2012, Willcox 2012). Agriculture first appeared here, at the Fertile Crescent, about 10-13.000 years ago (Willcox 2012). The process is often called the Neolithic Revolution, because of the rapid and fundamental changes in the culture and social structure experienced by these early populations. Within a couple of thousand years the first professions, the first stratified complex societies, and symbolic writing appeared (Wailes 1996, Gowdy and Krall 2013). Interestingly, this whole revolution originated in the change from mobile hunting and gathering to sedentary food producing. Our model captures elements of this change and shows how food production set the stage for SDL and specialisation.
Palaeolithic humans lived in small mobile groups, presumably had no hierarchy, no food storage, and no SDL (Boehm 2000, Simmons 2012, Gowdy and Krall 2013). In recent small-scale societies people are usually skilled in all subsistence related activities and use varying skills throughout a year. Everyone goes foraging and larger amounts of food are shared extensively (Kaplan et al. 1984, Marlowe 2010). In these societies fitness is related both to foraging success and food sharing (Kaplan and Hill 1985, Boehm 2001, Gurven and Von Rueden 2006). According to our model, in such a group, successful individuals are generalists, therefore selection creates a unimodal, symmetric distribution of the group’s genetic affinities (see Fig. 4.2). If there is a genetic trade-off, too, an extreme talent will bear a cost to its owner.

However, if food is storable and/or there is an opportunity for frequent trade, a chance appears to engage in SDL. We assume that this is exactly what happened during the Neolithic. If there is no pressure on individuals to learn all possible skills, they will get more or less specialised to tasks they have more talent for. Thus the cost/benefit ratio of task solving will decrease, and SDL will soon appear on the group level. SDL will lead to an increasing behavioural diversity, if specialised individuals differ from each other a great deal. Furthermore, if SDL and frequency dependence is present, selection will favour a multimodal distribution of the group’s genetic affinities (see Fig. 4.2). Thus, SDL will lead to a greater genetic diversity, too, on the group level. Therefore, the more collective tasks a society solves by SDL, the more diverse it will get both behaviourally and genetically. However, while phenotypic diversity appears very easily, genetic diversification needs special circumstances.

As we have mentioned in the Introduction, Nakahashi and Feldman (2014) studied a similar question to ours. They asked what the conditions are for DL to emerge if there are two types of resources that require different skills in a group living society with food sharing. For the sake of analytic tractability they assume asexual reproduction and individual learning. They have found that in certain circumstances permanent DL can emerge in a typical subsistence society, however, this is not in harmony with either the archaeological, or the anthropological record. As the only type of permanent DL observed in mobile small-scale societies is between the sexes, we interpret their model as best to capture the emergence of gender DL. Compared to them, we are mainly interested in the effect of the
Neolithic transition on SDL, as we believe that SDL only appeared together or following food production. Applying an individual based model, we incorporated sexual reproduction with recombination of parental traits and social learning (imitation), too. We also make a distinction between the phenotypic and genetic characteristics of individuals. Finally, we extended our model to cases with more than two tasks. Interestingly, despite all the differences, the two models find results in agreement with each other: a larger group size and an increased importance of learning a skill facilitates the emergence of DL, and thus seem to be crucial and robust requirements.

The above argument has implications not only for the Neolithic transition, but also for the more recent evolution of humans. In our model, the size of the group represents a major limit for genetic diversification. We have chosen a small basic group size that corresponds to the natural human social group size of about 150 people (Dunbar 1993). Presumably, Palaeolithic groups had sizes in this order of magnitude, where specialisation would indeed have been problematic. However, since the Neolithic, human groups are getting larger and larger. Therefore, we can easily assume that group size is no longer such a great limit on diversification. Moreover, with the IT revolution, the group during imitation or social learning is practically equal to the global population. Now that people’s mobility and connectedness has grown so drastically, SDL became global, and specialisation is ever more extreme. Thus, in the near future we may experience the greatest growth in human behavioural and genetic diversity.
Chapter 5

Conclusions

In this thesis I have presented three studies concerning social situations in structured populations. As each study had its own questions and methods, I have discussed each separately, however, a collective conclusion is also due. In the following I will summarise the general results we can draw from the work presented herein.

The emergence of cooperation and division of labour

Cooperative behaviour can be an evolutionary conundrum, primarily not because cooperation cannot be beneficial, but because its appearance and spread contradict the law of natural selection at first sight. DL, my other main focus, assumes mutual cooperation. However, for achieving a greater effectiveness in collective task solving, DL needs coordination, too. Here I have studied the emergence of cooperation and coordination in two different social settings.

In the first study, I have asked the question, whether rare cooperators can invade a population of defectors in different social dilemmas. There, I have found that, without viscosity, a rare mutant cooperator stands no chance against the resident defectors. However, in certain social dilemmas, such as the VD, viscosity can easily help even a single cooperator to locally outcompete defectors. I have also noted that such stable mixed equilibria are typical in the VD and in most threshold public goods games, but atypical in the NPD. If the VD is indeed a fairly good model of many social situations in nature, then its thorough study can help us understand both the invasion and the stable presence
of cooperators.

In the second study, my question was somewhat different. As here I already assumed cooperation, I only searched for the emergence of an optimal division of labour in a cooperative situation. Here, similarly to the previous case, I have found that in a homogeneous population performing one strategy there is a very low chance for the other strategy, and thus DL to appear. Nevertheless, if the population somehow reaches a state where both strategies are present in a certain, but not necessarily large quantity, a stable DL arises. Thus I have shown that in case of a rare strategy, the local mathematical analysis do not predict the global dynamics, because the population cannot reach the stable state, where branching can occur.

In summary, I have studied the spread of rare strategies in two quite different, but biologically relevant social situations. From these I can draw the shared conclusion that rare mutants, even if they bring the benefit of cooperation or division of labour, are easily outcompeted by resident types. Thus, for cooperation or DL to emerge, cooperators need some kind of disproportionately close relationship or a larger initial subpopulation. Also, the emergence of SDL is much easier in the model, where the benefit sharing resembles that of humans. Although the model operates with many simplifications, it can still add to the explanation, why SDL is much more frequent in humans than other animals.

Social division of labour and specialisation

In two of the presented studies I have examined the necessary conditions and possible evolutionary consequences of SDL. It is important to note here that, to my knowledge, the type of DL that we named SDL was first defined in Vásárhelyi et al. (2015). Permanent SDL is a remarkable social construction that is surprisingly rare in non-human animals. Therefore, to define and study its evolutionary background bears a special importance. In the second and third studies, I have examined models of different complexity, capturing various social situations. Nevertheless, two important questions were shared between these examinations: (i) What circumstances are needed for SDL to operate? (ii) What are the possible evolutionary consequences of SDL?

In the second study, we have found that SDL is easily maintained in relatively large
sections of the parameter space, if the trade-off between the costs of tasks is strong. However, SDL is much more typical in situations, where the benefit sharing is based on the members’ specific value to their group. In other words, if sharing is unequal with some kind of negative frequency dependence, SDL is easier to maintain. As for specialisation, we have seen that SDL (branching in the task choice probabilities) was always coupled with specialisation (cost optimisation). Fig. 3.4/B illustrates the process in detail: SDL is the first to occur, and when it is complete, specialisation follows. Thus, in this model the change in the task preferences was always followed by a change in performance. I note here, to avoid misunderstandings, that in this model both the task choice probabilities and the costs are heritable traits. Thus, here SDL always have a “genetic” basis, whereas in the third model SDL can occur throughout an individuals’s lifetime, too.

In the third study, I have considered a much more complex system. Here, initially all individuals had the capacity to solve all sorts of tasks for survival. Thus, the question was only whether SDL arose, thus rendering the group more effective, or did not. As I considered ecological circumstances, built in the fitness definitions, which either hindered or supported task specialisation, the emergence of SDL in the basic model turned out to be a mere question of the environment. However, as I enhanced the number of tasks dividable, phenotypic specialisation, and thus some kind of SDL was inevitable even if the environment was not supporting. Here, the relationship between SDL and genetic specialisation was not so strong, not obligatory. Because of imitation and the feedback loop of expertise, SDL appeared relatively easily in the producer case, and sometimes even in the subsistence case. However, genetic specialisation occurring only in the producer case was always incomplete, and required strong assortativity during imitation and mate choice. Finally, genetic specialisation did not occur without some, if not strong, negative frequency dependence. Thus the unequal share of resources enhanced specialisation here, too.

In summary, I have shown that SDL and specialisation can easily be coupled, and that both appear more easily if the share of benefits is unequal with some kind of negative frequency dependence.
Population size and population structure

So far, I have deliberately avoided talking much about the effect of the population’s size and structure on cooperation and SDL, because it deserves a section of its own. Let me separate these two factors though, as they can have quite different effects. The structure of a population is partly determined by its viscosity, whether it is divided into social groups or not, and if it is, then how large these groups are.

We have already seen that the invasion success of rare cooperators in the VD depended critically on population viscosity. Thus here the structure of the population was of crucial importance, even if the role of the social group size remained an open question to be studied later. The size of the population had less effect here. If the population was viscous, the size had no substantial effect whatsoever, but when the population was well-mixed, the size made some, if not a qualitative difference. It is important to see that the smaller the population, the larger probability the rare cooperators had for successful invasion. However, I repeat that this is only a quantitative effect, the invasion probability curves looked qualitatively the same (see also in Fig. 2.6). To summarise, in the VD the structure of the population turned out to be much more important than its size.

In the final part of the second study I asked what effect the population size had on the emergence of SDL. Although the populations considered here were smaller than natural populations are, I believe that the results are still instructive. I have found that the smaller a population, the more problematic the branching of trait values (genetic specialisation) became, reaching zero probability when $N_p = 150$ (see Fig. 3.6). This result is in absolute agreement with the results of the third study, as follows.

In the third study, I put an emphasis on the importance of the social group size. Here, although I assumed the groups to be part of a metapopulation, I did not include any sort of gene influx. Therefore, the social groups were also representing a closed reproductive pool. In this sense, by varying the size of the social group, I considered both the effect of the social group size and that of the reproductive pool. In this study, group size turned out to be a recurrent limit on task specialisation: with a growing number of tasks the necessary group size increased rapidly. Let me note here that this study contained a small subcase with a meta-population structure producing no qualitatively different
results. Interestingly thus, although group size is a quantitative factor, it can easily affect populations qualitatively.

To summarise, both the population’s size and structure can have a marked effect on cooperation and DL. While for the spread of rare mutants the viscosity of the population proved to be crucial, for specialisation (or branching) in the different models, the size was of equal importance. Although both viscosity and size are quantitative traits, in these studies they turned out to have qualitative effects, too.

**Summary of the main results**

In this thesis, I have discussed three studies on cooperation and division of labour. These studies, though differ both in their focus and complexity, are concerned with similar questions: How can cooperative behaviour and the coordination needed for DL appear? What circumstances are required for them to prevail? And what evolutionary consequences can we expect? In the following I will outline my main results obtained throughout the here presented work.

In the first study I have shown that

- Cooperation is more typical in $N$-person social dilemmas with sigmoid benefit functions, particularly in viscous populations.

- In agreement with the replicator dynamics of the VD, the invasion of cooperators in a well-mixed population is only possible if their initial concentration exceeds a critical threshold.

- In a viscous population, instead of the initial concentration, the initial number determines the success of invasion. A single mutant cooperator can invade with a high probability, since the local density of cooperators exceeds the critical threshold.

In the second study I have shown that

- SDL and specialisation can emerge in a population frequently facing a collective task, especially with a convex trade-off between the costs of different subtasks.

- SDL and specialisation are much more probable in the case of a benefit sharing system that takes into account the individual’s relative value to its group.
• Though the local mathematical analysis can predict the global behaviour, in the case of rare mutant strategies, the population never reaches the branching point, where specialisation could appear.

• Branching is seriously hindered by both social control and a small population size in the model.

Finally, in the third study I have shown that

• When food storage and/or trade becomes available, behavioural specialisation and large-scale division of labour is likely to appear.

• In the producer case both phenotypic and genetic specialisation is possible, if scarcer goods are more valuable.

• As the number of tasks increase, an ever larger group is necessary for specialisation.

• Although phenotypic specialisation is often present, genetic specialisation requires strong assortativity both during imitation and mate choice.
Acknowledgements

Throughout my PhD years I worked at the Department of Plant Systematics, Ecology and Theoretical Biology, ELTE, where I was always happy to spend my days. I am grateful that I had such wonderful colleagues, grateful not just for the inspiring conversations and the opportunity for teaching, but also for the lively lunches and serious ping-pong matches.

I also had the opportunity to work 3 months at the Centre for Behaviour and Evolution, Newcastle University, and 6 months at the KLI, Klosterneuburg. Both of these institutions, and my colleagues there, taught me a great deal about how science works, how Western European countries work, and how I should work. I am especially grateful to Daniel Nettle (CBE Newcastle) for supervising and supporting my experiment, and Yogi (Johannes Jaeger, KLI) for his honest(!) guidance and lots of useful advice.

Big thanks to Douglas Yu, Marco Archetti, Daniel Nettle and Yogi for reading and commenting the papers I cited here. Another big thank to Babszi Csapody and Mari Vecsenyi for correcting my English. Also thanks to Andris Hubai for helping with my figures (among other things).

And the most important is to thank István Scheuring. He started to supervise my work when I was doing my BSc degree. Many long years ago. Since then I learned an awful lot from him and also tried his patience sometimes, I am afraid. But never, for a moment, did I regret staying under his supervision, and I tell everything I need with this, I hope.

Finally, I would never have finished this thesis if Tamás and Palkusz did not let me neglect some of my motherly duties. It was a sacrifice on their part, and I am more than thankful for it.
Bibliography


Appendix A

Abbreviations and variables

C: Cooperator strategy
D: Defector Strategy
DL: Division of Labour
MC: Monte Carlo cycle
PD, NPD: Prisoner’s Dilemma Game, N-person Prisoner’s Dilemma Game
SD, NSD: Sowdrift Game, N-person Snowdrift Game
SDL: Social Division of Labour
VD: Volunteer’s Dilemma Game

α: the parameter defining the trade-off (Chapter 3 and 4)
β: a constant that converts expertise to benefit (Chapter 4)
γ: a constant regulating the strength of selection at imitations (Chapter 4)
δ: a constant scaling the strength of the frequency dependence (Chapter 4)
ε: a constant scaling the expertise function (Chapter 4)
λ: a constant scaling the steepness of the sigmoid benefit function (Chapter 2)
μ: a constant positioning the expertise function (Chapter 4)
ν: a constant positioning the maximum of the benefit function (Chapter 3)
ρ(x₀): invasion probability curve (Chapter 2)
σ²_i: the variance regulating the strength of assortativity at imitations (Chapter 4)
σ²_p: the variance regulating the precision of copying at imitations (Chapter 4)
σ²_R: the variance regulating the strength of assortativity at mate choice (Chapter 4)
σ²_G: the variance regulating the precision of genetic inheritance (Chapter 4)
\( \nu \): a parameter defining the range of mutations (Chapter 3)
\( \phi \): a parameter scaling the benefit function (Chapter 3)
\( \chi \): the magnitude of social control (Chapter 3)
\( \omega \): constant probability of dieing (Chapter 2)

\( B(i) \): group benefit with \( i \) cooperators (Chapter 2) and \( i \) playing \( T_1 \) (Chapter 3)
\( b \): benefit of cooperation, general notion (Chapter 2)
\( b_1(i), b_2(i) \): individual benefit for someone playing \( T_1 \) or \( T_2 \), respectively (Chapter 3)
\( b_{ij}(t) \): \( i \)'s net benefit from task \( j \) until round \( t \) (Chapter 4)
\( c \): cost of cooperation, general notion (Chapter 2)
\( c_1, c_2 \): variable costs of performing subtasks \( T_1 \) and \( T_2 \), respectively (Chapter 3)
\( d_{lm} \): Euclidean distance of the task choosing probabilities of \( l \) and \( m \) (Chapter 4)
\( E_{ij}(t) \): \( i \)'s expertise in task \( j \) in round \( t \) (Chapter 4)
\( E_0 \): a constant basic expertise (Chapter 4)
\( e_i \): elementary state of \( i \) cooperators in the population (Chapter 2)
\( f_j(t) \): frequency with which \( j \) was chosen till round \( t \) (Chapter 4)
\( G_i \): \( i \)'s genetics (Chapter 4)
\( G_{ij} \): \( i \)'s genetic affinity for task \( j \) (Chapter 4)
\( G_1^2, G_2^2 \): parental genetic values (Chapter 4)
\( h \): number of cooperators (Chapter 2)
\( h_{cr} \): inflection point of the benefit function (critical number of cooperators) (Chapter 2)
\( i \): number of cooperators in a group of \( N \) (Chapter 2)
\( k \): a constant cost (Chapter 3) or affinity (4)
\( L^2 \): torus size (Chapter 2)
\( M, M_0 \): neighbourhood, neighbourhood of an empty site (Chapter 2)
\( m \): intensity of mixing, number of mixing steps (Chapter 2)
\( N \): number of participants in a game (Chapter 2)
\( N_g \): group size (Chapter 2, Chapter 3) and 4)
\( N_T \): number of tasks (Chapter 4)
\( n_{ij} \): the number of rounds \( i \) already spent with task \( j \) (Chapter 4)
\( p \): probability in general (Chapter 2) and propensity of choosing \( T_1 \) (Chapter 3)
\( p^* \): socially controlled propensity of choosing \( T_1 \) (Chapter 3)

\( p_{ij} \): \( i \)'s probability of choosing \( j \) (Chapter 4)

\( p_j^X \), \( p_j^Y \): parental values for task choice probabilities (Chapter 4)

\( r \): the maximal difference between the cooperator and defector fitness (Chapter 2)

\( S = \left[ (1 + e^{-\lambda(i-h_{cr})})^{-1} \right] \): sigmoid benefit function (Chapter 2)

\( T_1, T_2 \): subtasks or strategies (Chapter 3) \( W_C, W_D, W_0 \): cooperator and defector fitness, basic fitness (Chapter 2)

\( W^+, W^0 \): higher and lower fitness, respectively, in the NVD (Chapter 2)

\( W_i^{(s)}, W_i^{(p)} \): \( i \)'s fitness in the subsistence and producer cases (Chapter 4)

\( x \)

\( x \): frequency of cooperators in the population (Chapter 2)

\( x = \{ q, c_A \} \): a strategy pair that characterises an individual (Chapter 3)

\( x^* \)

\( x^* \): fixed points of the replicators dynamics (Chapter 2)

\( x^* = \{ \frac{1}{2}, \frac{k}{\sqrt{2}} \} \): the singular point, where branching occurs (Chapter 3)

\( x_0 \): initial frequency of cooperators (Chapter 2)

\( x^{(0)} = \{ 0, k \} \), \( x^{(1)} = \{ 1, 0 \} \): stable fixed points at parameter end points (Chapter 3)

\( x_{cr} \): critical frequency of cooperators signifying a successful invasion (Chapter 2)

\( x_s, x_u \): stable and unstable inner fixed points of the replicator dynamics (Chapter 2)

\( x_u^*, x_u^{**} \): unstable fixed points between the end points and \( x^* \) (Chapter 3)
Appendix B
Supplementary for Chapter 2

NPD and VD games in infinite well-mixed populations

In the NPD, similar to the 2-person PD game, individuals follow either a cooperator (C) or a selfish defector (D) strategy in the NPD game too. Cooperators take on a costly act (with cost c) to provide a common good (b) which is equally divided among all group members, even those who did not pay a cost (defectors). We assume that the population is infinitely large and that all individuals form random groups of size $N_g$. The payoff of defectors ($P_D$) and cooperators ($P_C$) in groups having $i$ number of cooperators are

$$P_D = \frac{bi}{N_g}$$
$$P_C = \frac{bi}{N_g} - c = P_D - c$$

(S1)

respectively. Thus the average fitness of the strategies D and C are

$$W_D(x) = \sum_{i=0}^{N_g-1} f_{i,N_g-1}(x) \frac{bi}{N_g}$$
$$W_C(x) = \sum_{i=0}^{N_g-1} f_{i,N_g-1}(x) \left( \frac{b(i+1)}{N_g} - c \right),$$

(S2)

where

$$f_{i,N_g-1}(x) = \binom{N_g - 1}{i} x^i (1-x)^{N_g-1-i}$$

(S3)

is the probability of having $i$ cooperators among the other $N_g - 1$ group members if individuals are selected randomly from a population containing cooperators in fraction $x$. 
It is easy to see that \( W_C(x) = W_D(x) + b/N_g - c. \) Thus if \( c/b > 1/N_g \), defection is always better off. Naturally if \( c/b < 1/N_g \), the opposite is valid. The selection process can be described by the replicator equation (Taylor and Jonker 1978)

\[
\dot{x} = x(1-x) [W_C(x) - W_D(x)]. \tag{S4}
\]

In case of \( c/b > 1/N_g \), for every \( x \in [0,1] \) \( W_C(x) < W_D(x) \), and therefore \( x^* = 0 \) is the only stable fixed point of Equation (S4). If \( c/b < 1/N_g \), \( x^* = 1 \) is the only stable fixed point of the dynamics. There is no social dilemma in this case since \( C \) is the winner of the selection. However, it can be argued that usually \( b/c \) is smaller than the typical number of interacting individuals \( N_g \) (i.e. \( c/b > 1/N_g \)), and thus cooperation in general will go extinct (Hauert et al. 2006a).

Most of the biological examples of social dilemmas have a common feature: the benefit function increases monotonously in a non-linear and saturating manner. To model properly these situations, the public goods function in use, \( B(i) \), must meet the following criteria:

1. The public good increases monotonically with the number of cooperators in the group.
2. \( B(i) \) saturates (tends to a maximal value \( b \) as the number of cooperators \( i \) tends to the group size, \( i \to N_g \)).
3. At low \( i \), the increase in public goods production is accelerating, and after an inflection point \( (h_{cr}) \), the increase in public goods production becomes discounting (see Fig. 2.1 in the main text). (If \( h_{cr} \to 0 \), the benefit increases in a completely discounting manner, while if \( h_{cr} \to N_g \), the benefit is completely accelerating)

Thus, the average fitness of the strategies \( D \) and \( C \) are

\[
W_D(x) = \sum_{i=0}^{N_g-1} f_{i,N_g-1}(x)B(i) \\
W_C(x) = \sum_{i=0}^{N_g-1} f_{i,N_g-1}(x)(B(i+1) - c), \tag{S5}
\]
where $B(i)$ is a general public goods function satisfying the conditions listed above. It is biologically realistic to assume that the transition phase from accelerating to discounting is steep at the inflection point, and, thus, as a specific and easily tractable model, we can use a step function for $B(i)$ (Archetti 2009), which defines the generalised $N$-person Volunteer’s Dilemma Game (Archetti 2009, Diekmann 1985).

**Calculations concerning the initial steps of invasion by a rare mutant**

We considered a fraction of the grid with $5 \times 5$ points with one cooperator in the middle and defectors all around. If one of its neighbours dies, this cooperator can leave an offspring on the empty node with a probability proportional to its relative payoff. The probabilities of these events is calculated according to (2.3). We assumed that in every MC at most one of the $5 \times 5$ individuals dies, which is a good approximation if $\delta$ is small (in the corresponding simulations $\delta = 0.01$). Following this logic, one can calculate the probabilities of all possible transitions even for different “shapes” of cooperator patches within this $5 \times 5$ square (Fig. S1). Since the number of different possible patch shapes increases in a highly accelerating manner as the number of cooperators forming the patch increases, we did it only for the first three states (which means transitions $e_1 \rightarrow e_2$, $e_1 \rightarrow e_0$, $e_2 \rightarrow e_3$, $e_2 \rightarrow e_1$).

According to the law of total probability, one can calculate the probabilities in question as follows:

$$
p{1 \rightarrow \ldots \rightarrow 3} = p{1 \rightarrow 2}p{2 \rightarrow \ldots \rightarrow 3}
$$

$$
p{2 \rightarrow \ldots \rightarrow 3} = p{2 \rightarrow 1}p{1 \rightarrow \ldots \rightarrow 3} + p{2 \rightarrow 3}
$$

$$
\Rightarrow
$$

$$
p{1 \rightarrow \ldots \rightarrow 3} = (p{1 \rightarrow 2}p{2 \rightarrow 3})/ (1 - p{1 \rightarrow 2}p{2 \rightarrow 1}), \quad (S6)
$$

where the expression $p{i \rightarrow \ldots \rightarrow i + 1}$ means a transition with any number of
intermediate states before reaching state $e_i$, while $p\{i \rightarrow i + 1\}$ means a transition with no intermediate state. To compute $p\{i \rightarrow i + 1\}$ we have to study transition probabilities for all topologically different cases and compute the weighted sum of these probabilities (see Fig. S1 for topologies and notations). Here we only show the calculation of one probability value ($p\{e_2 \rightarrow e_{3C}\}$), but all the others can be calculated similarly.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure_s1.png}
\caption{Possible topologies of states $e_1$, $e_2$ and $e_3$. Black squares are cooperators and grey squares are positions where these cooperators can leave offspring to. The width of the arrows represents the number of possible ways how an $e_i$ state develops into a given $e_{i+1}$ state (thin line one way, thick line two ways).}
\end{figure}

The two topologically different $e_2$ states ($e_{2A}$ and $e_{2B}$) develop from $e_1$ with equal probability, but this is not true in case of the five different $e_3$ states. For example $e_{3C}$ (in the middle of the bottom of Fig. S1) can develop from both $e_2$ states, but with different probabilities. These probabilities depend on several parameters:

- the number of those neighbouring empty grid points, from which the given configuration can evolve (for example the $e_{2B} \rightarrow e_{3C}$ transition can not happen if a diagonal neighbour dies)

- the number of cooperators neighbouring each of these points
the fitness values of all neighbours of these points

Let us denote the probabilities of transitions between these topologically different $e_i$ states by $p_e$. Using notations defined in (2.2), the probability of transition $e_2 \rightarrow e_{3C}$ (denoted by $p_e\{2 \rightarrow 3C\}$) is

$$p_e\{2 \rightarrow 3C\} = \frac{1}{2} \left( \frac{4}{25} \frac{2W_C^+}{2W_C^+ + W_D^+ + 5W_D^0} + \frac{2}{25} \frac{2W_C^+}{2W_C^+ + W_D^+ + 5W_D^0} \right),$$

(S7)

where the $\frac{1}{2}$ is needed because of the two different $e_2$ states, $\frac{4}{25}$ and $\frac{2}{25}$ are the probabilities of having those empty neighbouring positions occupying which $e_{3C}$ would evolve from $e_{2A}$ and $e_{2B}$ respectively and the two $\frac{2W_C^+}{2W_C^+ + W_D^+ + 5W_D^0}$ expressions are the probabilities of leaving an offspring with cooperator strategy to these empty places (see also Models and Methods). (Note that these ratios are usually not equal.)

Following the logic of (S7), one can calculate the probabilities of all $\{e_{im} \rightarrow e_{jn}\}$ transitions, where $i = \{1; 2\}$, $j = \{2; 3\}$, $m \in M := \{A; B\}$ and $n \in N := \{A; B; C; D; E\}$. The $\{e_i \rightarrow e_j\}$ probabilities emerge as the sum of these values. For example $p_e\{2 \rightarrow 3\} = \sum_{m \in M} \sum_{n \in N} p\{2m \rightarrow 3n\}$

The last step to get the probabilities $p\{1 \rightarrow \ldots \rightarrow 3\}$, $p\{2 \rightarrow \ldots \rightarrow 3\}$ and $p\{1 \rightarrow \ldots \rightarrow 3\}$ from Equation S6 is to calculate the $p\{i \rightarrow i+1\}$ probabilities using the $p_e$ values. For example

$$p\{i \rightarrow i+1\} = \frac{p_e\{i \rightarrow i+1\}}{p_e\{i \rightarrow i+1\} + p_e\{i \rightarrow i-1\}}.$$  

Using numerical values counted the way we presented above, we get the following probabilities for the basic parameter set with $h_{cr} = 2$:
\[
p\{1 \to 2\} = 0.497246 \\
p\{1 \to 0\} = 0.502754 \\
p\{2 \to 3\} = 0.519639 \\
p\{2 \to 1\} = 0.480361 \\
\]

\[
p\{1 \to \ldots \to 3\} = 0.317291 \\
p\{2 \to \ldots \to 3\} = 0.682709 \\
\]

Note that in neutral case, the probabilities of all of the \(p\{i \to j\}\) transitions would be 0.5. In the neutral case, \(p\{1 \to \ldots \to 3\} = 1/3\) and \(p\{2 \to \ldots \to 3\} = 2/3\).
Appendix C

The analytic model of Chapter 3

The analytic model of social division of labour

We discuss the 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

$$
\Delta b(i) = b_1(i + 1) - b_2(i) 
$$

(S8)

is the benefit advantage (or disadvantage) of playing strategy $T_1$ over strategy $T_2$ for an individual, provided that the rest of the group have $i$ members playing $T_1$. It is easy to see that

$$
\Delta b(i) = -\Delta b(N_g - 1 - i),
$$

(S9)

which we will use later on.

Payoffs

The two independent heritable traits of individuals are $p \in [0, 1]$, the probability of choosing subtask $T_1$, and $c_1 \in [0, k]$, the cost assigned to this task. The expected payoff values for an individual playing both strategies are:

$$
w_1(p, c_1) = \sum_{i=0}^{N_g-1} g_{i,N_g-1}(p)b_1(i + 1) - c_1
$$

$$
w_2(p, c_1) = \sum_{i=0}^{N_g-1} g_{i,N_g-1}(p)b_2(i) - (k^{\alpha} - c_1^{\alpha})^{\frac{1}{\alpha}},
$$

(S10)
where \( p \) denotes the average \( p \) value of all the other population members, i.e. the probability of a randomly chosen individual playing \( T_1 \). The value for \( c_2 \) was substituted. The probability of having \( i \) individuals playing \( T_1 \) among the other \( N_g - 1 \) group members is given by the binomial distribution

\[
g_{i,N_g-1}(\bar{p}) = \binom{N_g - 1}{i} \bar{p}^i (1 - \bar{p})^{N_g - 1 - i} \tag{S11}
\]

obeying the symmetry relation

\[
g_{i,N_g-1}(\bar{p}) = g_{N_g-i,N_g-1}(1 - \bar{p}). \tag{S12}
\]

Finally, the expected payoff for an individual with trait values \( \{p, c_1\} \) is

\[
\overline{w}(p, c_1, \bar{p}) = pw_1(\bar{p}, c_1) + (1 - p)w_2(\bar{p}, c_1). \tag{S13}
\]

**Cost optimisation**

The sub-problem of cost-minimisation of an individual with a fixed \( p \) value can be considered separately, because it depends only on the \( p \) value of the individual and contributes to the payoff additively. Therefore, our first interest is to find the optimal \( c_1 \) cost strategy that minimises the average cost

\[
\overline{c}(c_1) = pc_1 + (1 - p)c_2 = pc_1 + (1 - p)(k^\alpha - c_1^\alpha)^{\frac{1}{\alpha}} \tag{S14}
\]

for a fixed \( p \).

For \( \alpha = 1 \) it reduces to

\[
\overline{c}(c_1) = pc_1 + (1 - p)(k - c_1) = k(1 - p) + (2p - 1)c_1, \tag{S15}
\]

which is monotonous in \( c_1 \). For \( p < \frac{1}{2} \) it means that cost minimisation results in \( c_1 = k \) (full specialisation for task \( T_2 \)), while \( p > \frac{1}{2} \) makes \( c_1 = 0 \) (full specialisation for task \( T_1 \)) the optimal strategy. For \( p = \frac{1}{2} \), \( c_1 \) is a neutral trait, which does not affect the average cost.

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

\[
\frac{d\overline{c}(c_1)}{dc_1} = p - (1 - p)c_1^{\alpha-1}(k^\alpha - c_1^\alpha)^{(1-\alpha)/\alpha} = 0, \tag{S16}
\]
leading to
\[ c_1^{\text{ext}}(p) = k \left[ \left( \frac{p}{1 - p} \right)^{\frac{\alpha}{\alpha - 1}} + 1 \right]^{-\frac{1}{\alpha}} \] (S17)

extreme \( c_1 \) value. This function decreases for \( \alpha < 1 \) and increases for \( \alpha > 1 \). Relatedly, as it can be seen in Figure 3.1, 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. (S17), while for \( \alpha \geq 1 \), it is at one of the end-points according to
\[ c_1^{\text{min}}(p) = \begin{cases} 0, & \text{for } p > \frac{1}{2} \\ k, & \text{for } p < \frac{1}{2}. \end{cases} \] (S18)

In the latter case the two end points are equally optimal for \( p = \frac{1}{2} \).

The cost optimisation problem is trivial: \( p = 0 \), or 1, when the individual always choose the same subtask. Then it is optimal to fully specialise to the same subtask, leading to zero cost, and pessimal to specialise to the other subtask, leading to the maximal cost \( k \). If \( p \) changes from 0 to 1, the optimal \( c_1 \) changes from \( k \) to 0. This transition is continuous for \( \alpha < 1 \), but abrupt at \( p = \frac{1}{2} \) for \( \alpha > 1 \). It can be seen that the optimal \( \bar{c} \) increases with \( p \) until \( p = \frac{1}{2} \) 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 \( x: x = \{p, c_1\} \) 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
\[ W(x', x) = \overline{w}(p', c_1', p) - \overline{w}(p, c_1, p), \] (S19)
which is 0 by definition for $x' = x$. (Since the mutant is rare, it does not modify $p$ in the leading term.)

Note that the fitness is linear in $p'$, but is not so in $p$. Linearity in a rare trait is a common feature of game theoretic models with mixed strategies (Hofbauer and Sigmund 1998), leading to non-standard behaviour relative to standard adaptive dynamics (Meszéna et al. 2001). On the other hand, the non-linearity in $p$ 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

$$\frac{\partial W(x', x)}{\partial x'_i} \bigg|_{x' = x} = 0,$$

(S20)
i.e. if

$$\frac{\partial W(x', x)}{\partial p'} \bigg|_{x' = x} = w_1(p, c_1) - w_2(p, c_1) = 0,$$

(S21)
and if

$$\frac{\partial W(x', x)}{\partial c'_1} \bigg|_{x' = x} = -\frac{d\sigma(c_1)}{d c_1} = 0.$$

(S22)
We have already solved Eq. (S22) by Eq. (S17) for $\alpha \neq 1$. Substituting it into (S21), $p$ is determined by

$$G(p) + H(p) = 0$$

(S23)
at the singular point, where

$$G(p) = \sum_{i=0}^{N_g-1} g_{i, N_g-1}(p) \Delta b(i)$$

(S24)
is the expected benefit advantage for the focal individual to play A, instead of B, and

$$H(p) = -\frac{d\sigma(c_1^{\text{extr}}(p))}{d p} = k \left[ \left( \frac{p}{1-p} \right)^{\frac{1}{1-\alpha}} - 1 \right] \left[ \left( \frac{p}{1-p} \right)^{\frac{\alpha}{1-\alpha}} + 1 \right]^{-\frac{1}{\alpha}}$$

(S25)
is the contribution from the cost term, valid for $\alpha \neq 1$.

It follows from the symmetry of $\Delta b(i)$ and $\Delta g_{i, N_g-1}(p)$ (see (S9) and (S12)) that $G(p) = -G(1-p)$, so $G(1/2) = 0$ for model I) and model II) as well. One can see that $H(p)$ is also zero for $p = 1/2$; that is (S21) is satisfied for $p = 1/2$ at $\alpha \neq 1$. 

109
The case $\alpha = 1$ requires separate consideration, as the internal extremum of the average cost $c_1$ does not exist in this case. Recall however, that cost is a neutral trait for $\alpha = 1$, $p = \frac{1}{2}$. That is, the singular point is replaced by a singular line at $p = \frac{1}{2}$. This situation is usually neglected in adaptive dynamics.

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

$$
\lim_{p \to 0} H(p) = -\theta k
$$

(S26)

$$
\lim_{p \to 1} H(p) = \theta k,
$$

(S27)

where

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

(S28)

that is, for $\alpha = 1$ $H(p) \equiv 0$. Consequently, $p^* = 1/2$ is the only solution of (S23) 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

$$
G(1) + H(1) > 0
$$

(S29)

$$
\left. \frac{d [G(p) + H(p)]}{dp} \right|_{p=1/2} < 0,
$$

(S30)

then at least one additional pair of singular points exists besides $x^*$ (Fig. S2). Denote these points by $x_u^* = \{p_u^*, c_u^*\}$, where $p_u^* > 1/2$ and $x_u^{**} = \{p_u^{**}, c_u^{**}\}$, where $p_u^{**} = 1 - p_u^* < 1/2$, 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 $T_1$ and this subtask is cost-free. Now we study the invasion of a mutant in the neighbourhood of $x^{(1)}$ by asking whether there are $x' = \{p', c'_1\}$ mutants with higher fitness than the resident one, that is whether

$$W(x', x^{(1)}) \approx \frac{\partial W(x', x)}{\partial p'} \bigg|_{x'=x^{(1)}} \delta p + \frac{\partial W(x', x)}{\partial c'_1} \bigg|_{x'=x^{(1)}} \delta c_1 \quad (S31)$$

can be positive, or not, where $\delta p = p' - p^{(1)} = p' - 1 \leq 0$ and $\delta c_1 = c'_1 - c^{(1)}_1 = c'_1 \geq 0$ in such a way that at least one of these terms differs from zero. It is clear that the partial derivatives become

$$\frac{\partial W(x', x)}{\partial p'} \bigg|_{x'=x^{(1)}} = G(1) + k, \quad (S32)$$
$$\frac{\partial W(x', x)}{\partial c'_1} \bigg|_{x'=x^{(1)}} = -1. \quad (S33)$$

The second term in (S31) is always negative ($-1 \times \delta c_1 < 0$). Thus $W(x', x^{(1)})$ can be positive only if $G(1) + k$ is negative (since $\delta p < 0$), which is the condition for the invasion.
of a mutant (if \( \delta c_1 = 0 \) then \( p' < 1 \) 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:

\[
G(1) + k = \begin{cases} 
[B(N_g) - B(N_g - 1)]/N_g + k, \text{ for model I),} \\
B(N_g)/N_g - B(N_g - 1) + k, \text{ for model II) }
\end{cases}
\]  

(S34)

Considering model I) it is clear that \( B(N_g) - B(N_g - 1) < 0 \), thus (S34) can be negative. However, if \( N_g \) 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, \( B(N_g)/N_g - B(N_g - 1) \) is negative, but this value is smaller than in the previous case, thus \( G(1) + k \) can easily remain negative if \( N_g \) is large and \( B(N_g - 1) > k \).

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 (3.2) has a zero-or-infinite derivative at \( c_1 = 0, \) or 1, for \( \alpha \neq 1 \). That is, if mutation steps are uniformly small, when parametrised by \( c_1 \), they become non-uniform when re-parametrised by \( c_2 \). From a biological point of view, the natural assumption is that mutations are uniformly small in \( c_1 \) in the vicinity of \( c_1 = 0 \) and uniformly small in \( c_2 \) in the vicinity of \( c_2 = 0 \). Thus \( c_2 \), 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 \( p \) 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 \( p' < 1 \) and \( c'_1 \geq 0 \) 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 \( c'_1 \) is

\[
\frac{\partial W(x'', x')}{\partial c'_1} \bigg|_{x'' = x'} = \begin{cases} 
-p' + (1 - p')O((c'_1)^{\alpha-1}), \text{ for } \alpha > 1, \\
1 - 2p', \text{ for } \alpha = 1, \\
-p' + (1 - p')/O((c'_1)^{\alpha-1}), \text{ for } \alpha < 1
\end{cases}
\]  

(S35)

(cf. (S16)). If \( \alpha < 1 \) then this derivative is a large positive quantity for very small \( c_1 \).
Therefore $c_1'' > 0$ easily spreads, and the invasion process continues to the inner singular point. Recall that the cost optimum for a given $p$ is internal for $\alpha < 1$. Consequently, the “best” direction for a small mutation follows the direction of the curve $c_1^{\text{extr}}(p)$. That is, we have an essentially one-dimensional dynamics along this curve; direction is determined by the sign of $-p' \approx -1$, thus invasion of $c_1'' > 0$ is not supported and invasion events stop in $x^m$ where $p^m < p^{m-1}$, $c_1^m = c_1^{m-1} = 0$. Because the cost optimum is always at one of the end points of the $c_1$ scale, if the pure strategy is invaded, then evolution proceeds along the edge $c_1 = 0$ of the strategy space. Direction of this one-dimensional strategy is determined by the sign of $G(p) + k$, 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 $p > \frac{1}{2}$, i.e. before the cost optimum switches to the other extreme.

**Convergence and branching of the singular point $x^*$**

The question is the $x^* = \{p^*, c_1^*\}$ singular point’s convergence stability and the conditions of branching. The convergence of a singular point means that in its neighbourhood, 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, 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

$$H_{i,j} = \frac{\partial^2 W(x', x)}{\partial x'_i \partial x'_j} \bigg|_{x'=x} \quad \text{(S36)}$$

$$Q_{i,j} = \frac{\partial^2 W(x', x)}{\partial x'_i \partial x_j} \bigg|_{x'=x}, \quad \text{(S37)}$$

and $x_i$ denotes the coordinates of the $x$ vector. If $J$ 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 $H$ 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 $J$ matrix becomes

$$\begin{pmatrix}
Q_{1,1} & H_{1,2} \\
H_{2,1} & H_{2,2}
\end{pmatrix} \quad \text{(S38)}$$
We have seen that if $B(i)$ is symmetrical then $\{p^*,c_1^*\} = \{\frac{1}{2}, \frac{k}{\sqrt{2}}\}$ becomes a singular point. Then the $J$ matrix becomes:

$$
\begin{pmatrix}
\Delta w(p^*) & -2 \\
-2 & \frac{\alpha-1}{c_1^*}
\end{pmatrix},
$$

(S39)

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

$$
\Delta w(p^*) = \left. \frac{\partial[w_1(p,c_1) - w_2(p,c_1)]}{\partial p} \right|_{p=p^*,c_1=c_1^*}.
$$

(S40)

After the derivation we get

$$
\Delta w(p^*) = \sum_{i=0}^{N_g-1} \binom{N_g-1}{i} p^{*i-1} (1-p^*)^{N_g-2-i} [i - (N_g-1)p^*] \Delta b(i).
$$

(S41)

Using $p^* = 1/2$, the symmetry of $B(i)$ and Equation (S24) we get

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

(S42)

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

The $\{p^*,c_1^*\} = \{\frac{1}{2}, \frac{k}{\sqrt{2}}\}$ point is convergence stable (eigenvalues of (S39) are negative) if

$$
\Delta w(1/2) + \frac{\alpha-1}{c_1^*} < 0
$$

$$
\Delta w(1/2) \frac{\alpha-1}{c_1^*} - 4 > 0.
$$

(S43)

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

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

(S44)

We note here that $\Delta w(1/2)$ is greater for model I) than for model II) at the same $B(i)$, thus the above criterion can be satisfied with a smaller $\alpha$ (stronger trade-off) in model I) than in model II). Similarly, $\Delta w(1/2)$ is greater if the maximal value of $B(i)$ 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:

\[
\begin{pmatrix}
0 & -2 \\
-2 & \frac{\alpha-1}{c_1}
\end{pmatrix},
\]  
(S45)

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 summarise: the \( \{p^*, c_1^*\} = \{\frac{1}{2}, \frac{k}{\sqrt{2}}\} \) singular point can only be convergence stable if \( \alpha < 1 \), which is fulfilled if condition (S44) is true. In this case the singular point is a branching point too. If \( \alpha \geq 1 \), both \( J \) and \( H \) will be indefinite. In this case, the convergence of the singular point depends on the covariance matrix of the mutants (Leimar 2009).
Appendix D
Supplementary for Chapter 4

Figure S3: Summary of the results in both settings with $N = 1000$ and stronger trade-off, $\alpha = 0.75$. The figure shows the task choice probabilities (top row) and the distribution of the genetics (bottom row) for one task throughout all generations. The top right plot show the fitness of individuals as a function of their average life-long task choice probabilities. The bottom right plot show these average probabilities as function of the genetic affinity assigned to the same task. The greenish colours (A columns) represent the subsistence, the reddish (B columns) the producer case. For other parameters see table 4.1.
Figure S4: Summary of the results with varying trade-off and frequency dependence in the producer case. The figure shows the task choice probabilities (top row) and the distribution of the genetics (bottom row) for one task throughout all generations. Here frequency dependence is much weaker ($\delta = 0.25$ and $\delta = 0.4$) than in the basic parameter set ($\delta = 0.01$). For other parameters see table 4.1.
Figure S5: Phenotypic and genetic profiles in the $T = 3$ case with varying trade-off and increased task learning speed ($\epsilon = 0.2$). $N = 500$ and $\gamma = 5$. In the top row simplexes capture the probability distribution between the three tasks, in the bottom row the relative genetic value distribution between the three tasks. Each point corresponds to an individual, and plots contain all group members. For further parameters, see table 4.1.
Figure S6: Phenotypic (top) and genetic (bottom) profiles in the $T = 4$ case with stronger trade-off ($\alpha = 0.75$). The figure captures tetrahedron nets, where each triangle is a simplex calculated from the three noted tasks, similarly to the simplexes on fig. 4.3. Thus, the faces of the tetrahedrons show projections of the 3D mass of points. Each point corresponds to an individual, and plots contain all group members. Reddish colours represent the producer, greenish the subsistence case. For further parameters, see table 4.1.
Figure S7: Summary of the results with four tasks, overlapping genetics and stronger trade-off ($\alpha = 0.75$) in the producer case. The figure shows the task choice probabilities (left and middle columns) and the distribution of the genetics (right column) for one task throughout all generations. For further parameters, see table 4.1.
Figure S8: Summary of the results with four tasks, overlapping genetics and weaker trade-off ($\alpha = 1.25$) in the producer case. The figure shows the task choice probabilities (left and middle columns) and the distribution of the genetics (right column) for one task throughout all generations. For further parameters, see table 4.1.
Major Evolutionary Transitions are often characterised by cooperation and division of labour on a given level of organisation which leads to the emergence of a higher level of evolutionary unit. In my thesis I attempt to answer a number of untackled questions regarding the emergence and evolutionary role of cooperation and division of labour in structured populations, with special emphasis on human social groups. I consider three problems, as follows.

In the first study I ask how rare mutant cooperators can invade a population of cheaters in different social dilemmas. With a cellular automata model I show that in the Volunteer’s Dilemma, the coexistence of cooperators and defectors is typical, however, the invasion of cooperators in a well-mixed population is only possible if their initial concentration exceeds a critical threshold. In a viscous population, however, even a single mutant cooperator can invade with a high probability.

In the second study I analyse a strategic model that studies the emergence of division of labour and specialisation. I show that division of labour and specialisation can emerge in a population, where small groups frequently face a collective task. Yet, in the case of rare mutant strategies, the population cannot leave its homogeneous state. Specialisation is also seriously hindered by both social control and a small population size in the model.

In the third study I discuss an individual based model that considers the connections between the appearance of food production, human social division of labour and behavioural diversity. I show that when food storage and/or trade becomes available, behavioural specialisation and large-scale division of labour is likely to appear. Although phenotypic specialisation is often present, genetic specialisation requires strong assortativity both during imitation and mate choice.
Összefoglalás

Az evolúció nagy átmeneteit gyakran jellemzi olyan együttműködés és munkamegosztás egy adott szerveződési szinten, amely egy magasabb szintű evolúciós egység megjelenéséhez vezet. Doktori diszsertációmban még megválaszolatlan kérdésekre kerestem a választ az együttműködés és a munkamegosztás megjelenésével, evolúciós következményeivel kapcsolatban. A dolgozatban strukturált populációkkal foglalkozom, különös tekintettel az emberi társadalmakra. A dolgozat három problémakörrel foglalkozik.

Az első tanulmányban azt vizsgálok, hogy a ritka mutáns kooperátorok hogyan terjednek a tiszta csaló populációkban különböző szociális dilemmák esetén. Egy sejtautomata modellel megmutatom, hogy az Önkéntesek Dilemmájában megvalósul a kooperátorok és a csalók együttélése, ugyanakkor a kooperátorok terjedése a jólkevert modellben csak akkor lehetséges, ha a kezdeti gyakoriságuk magasabb egy kritikus küszöbértéknel. A viszkózus esetben ezzel szemben egyetlen kooperátor is nagy esélyvel képes terjedni.

A második tanulmányban egy a munkamegosztás örökletes viselkedési különbségekre való hatását vizsgáló stratégiai modellt elemeztem. Megmutattam, hogy a munkamegosztás és a specializáció könnyen megjelenhet egy olyan populációban, amelynek kis csoportjai ismételten szembesülnek egy kooperatív feladattal. Ha azonban az egyik stratégia ritka mutánsként van jelen, a populáció sohasem válik szét specializált felekre. A szétválaszt mind a szociális kontroll, mind a kis populációméret erősen gátolja.

A harmadik tanulmányban egy olyan egyed alapú modellt tárgyalok, amely az élelmiszertermelés, a szociális munkamegosztás és a viselkedési sokféleség kapcsolatát vizsgálja. Megmutatom, hogy a táplálék raktározását és/vagy a keresedelmet követően könnyen megjelenhet a specializáció és a széles munkamegosztás is. Míg a fenotípusos specializáció nagyon gyakran létrejön, a genetikai specializációhoz erős asszortativitásra van szükség mind az imitáció, mind a párválasztás során.
ADATLAP

a doktori értekezés nyilvánosságra hozatalához*

I. A DOKTORI ÉRTEKEZÉS ADATAI

A szerző neve: Vásárhelyi Zsóka Anna

MTMT-azonosító: 10052969

A doktori értekezés címe és alcíme: The evolution of cooperation and division of labour in structured populations – Three individual based models

DOI-azonosító: 10.15476/ELTE.2018.040

A doktori iskola neve: Biológia Doktori Iskola

A doktori iskolán belüli doktori program neve: Elméleti és evolúcióbiológia

A témavezető neve és tudományos fokozata: Scheuring István, tudományos tanácsadó

A témavezető munkahelye: MTA OK Evolúciós Rendszerek Kutatócsoport

II. NYILATKOZATOK

1. A doktori értekezés szerzőjeként

a) hozzájárulok, hogy a doktori fokozat megszerzését követően a doktori értekezésem és a tézisek nyilvánosságra kerüljenek az ELTE Digitális Intézményi Tudásvállalatban. Felhatalmazom a Természettudományi kar Dékáni Hivatal Doktori, Habilitációs és Nemzetközi Ügyek Csoportjának ügyintézőjét, hogy az értekezést és a téziseket feltöltse az ELTE Digitális Intézményi Tudásvállalatba, és ennek során kitöltse a feltöltéshez szükséges nyilatkozatokat.

b) kérem, hogy a mellékelt kérelemben részletezett szabadalmi, illetőleg oltalmi bejelentés közzétételét a doktori értekezést ne bocsássák nyilvánosságra az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudásvállalatban;

c) kérem, hogy a nemzetbiztonsági okból minősített adatot tartalmazó doktori értekezést a minősítés (dátum)-ig tartó időtartama alatt ne bocsássák nyilvánosságra az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudásvállalatban;

d) kérem, hogy a mű kiadására vonatkozó mellékelt kiadó szerződésre tekintettel a doktori értekezést a könyv megjelenéséig ne bocsássák nyilvánosságra az Egyetemi Könyvtárban, és az ELTE Digitális Intézményi Tudásvállalatban csak a könyv bibliográfiai adatait tegyék közzé. Ha a könyv a fokozatátadást követő egy évig nem jelenik meg, hozzájárulok, hogy a doktori értekezésem és a tézisek nyilvánosságra kerüljenek az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudásvállalatban.

2. A doktori értekezés szerzőjeként kijelentem, hogy

a) az ELTE Digitális Intézményi Tudásvállalatra feltöltendő doktori értekezés és a tézisek saját eredeti, önálló szellemi munkám és legjobb tudomásom szerint nem sért sem véle senki szerzői jogait;

b) a doktori értekezés és a tézisek nyomtatott változatai és az elektronikus adathordozón benyújtott tartalmak (szöveg és ábrák) mindenben megegyeznek.

3. A doktori értekezés szerzőjeként hozzájárulok a doktori értekezés és a tézisek szövegének plágiumkereső adatbázisba helyezéséhez és plágiumellenőrző vizsgálatok lefuttatásához.


a doktori értekezés szerzőjének aláírása

*ELTE SZMSZ SZMR 12. sz. melléklet