

## Article

# Social Learning versus Individual Learning in the Division of Labour

Moein Khajehnejad , Julian García  and Bernd Meyer \* 

Department of Data Science and Artificial Intelligence, Monash University, Clayton, VIC 3168, Australia

\* Correspondence: bernd.meyer@monash.edu

**Simple Summary:** Division of labour is a crucial characteristic of social organisations such as insect colonies and is a key feature in their well-known survival and efficacy. The presence of “laziness”, or inactivity is a widely debated phenomenon that has been observed in some colonies and is puzzling because it goes against the idea that a division of labour would lead to greater efficiency and effectiveness. Inactivity has been previously explained as a by-product of social learning, which is a fundamental type of behavioural adaptation in these colonies. However, this explanation is limited because it is still unclear if social learning governs aspects of colony life. This study explores how inactivity can also emerge similarly from an individual learning paradigm, which is a firmly established paradigm of behaviour learning in insect colonies. Using individual-based simulations backed up by mathematical analysis, the study finds that individual learning can induce the same behavioural patterns as social learning. This is important for understanding the collective behaviour of social insects. The insight that both modes of learning can lead to the same patterns of behaviour opens up new ways of approaching the study of emergent patterns of collective behaviour in a more generalised manner.



**Citation:** Khajehnejad, M.; García, J.; Meyer, B. Social Learning versus Individual Learning in Division of Labour. *Biology* **2023**, *12*, 740. <https://doi.org/10.3390/biology12050740>

Academic Editor: Antonio Carvajal-Rodríguez

Received: 16 February 2023

Revised: 15 May 2023

Accepted: 17 May 2023

Published: 19 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Division of labour, or the differentiation of the individuals in a collective across tasks, is a fundamental aspect of social organisations, such as social insect colonies. It allows for efficient resource use and improves the chances of survival for the entire collective. The emergence of large inactive groups of individuals in insect colonies sometimes referred to as *laziness*, has been a puzzling and hotly debated division-of-labour phenomenon in recent years that is counter to the intuitive notion of effectiveness. It has previously been shown that inactivity can be explained as a by-product of social learning without the need to invoke an adaptive function. While highlighting an interesting and important possibility, this explanation is limited because it is not yet clear whether the relevant aspects of colony life are governed by social learning. In this paper, we explore the two fundamental types of behavioural adaptation that can lead to a division of labour, *individual learning* and *social learning*. We find that inactivity can just as well emerge from individual learning alone. We compare the behavioural dynamics in various environmental settings under the social and individual learning assumptions, respectively. We present individual-based simulations backed up by analytic theory, focusing on adaptive dynamics for the social paradigm and cross-learning for the individual paradigm. We find that individual learning can induce the same behavioural patterns previously observed for social learning. This is important for the study of the collective behaviour of social insects because individual learning is a firmly established paradigm of behaviour learning in their colonies. Beyond the study of inactivity, in particular, the insight that both modes of learning can lead to the same patterns of behaviour opens new pathways to approach the study of emergent patterns of collective behaviour from a more generalised perspective.

**Keywords:** division of labour; social learning; individual learning; evolutionary game theory; adaptive dynamics; cross-learning

## 1. Introduction

Division of labour is fundamental to the functioning of social organisms and has been central to their study for decades [1]. The separation of tasks among different individuals or groups within a collective allows for the efficient use of resources and increases the chances of survival for the collective as a whole [2–4]. Studies have shown that division of labour is prevalent in many socially living organisms, such as ants, bees, termites, and even some mammals [5–9]. Social insect colonies are well known for their intricate organisation and their ability to handle a wide range of tasks simultaneously, including foraging, colony defence, nest construction, temperature regulation, and caring for offspring [6,10]. The colony's ability to effectively allocate its workforce to these different tasks, adapting to changes in both external conditions and internal needs, is often cited as a key to their ecological success [11–14]. Understanding the underlying mechanisms of division of labour is fundamental to understanding these social organisations and the emergence of complex social systems in general.

It is well established that both developmental and genetic factors significantly influence the division of labour [15,16]. Additionally, studies have shown that faster and self-organised mechanisms for division of labour exist within colonies, enabling them to rapidly and adaptably respond to shifts in task requirements. Environmental factors or internal shifts within the colony may be responsible for these changes [14,17]. The fast changes in labour division arise from a combination of factors including workforce distribution, interaction structures, and environmental influences [18–21]. Empirical research has also emphasised the significance of social context and interactions in shaping the task preferences of individuals [22,23]. Individuals generally lack knowledge of the overall state of the colony, thus their behavioural decisions rely on the local information that is readily available to them [24,25]. Interactions among colony members can offer valuable insight into the colony's condition and act as cues for behaviour, as well as a means for social learning [26,27]. Information acquired through local interactions with other individuals and the environment can often indicate the global state of the colony.

While not frequently discussed in connection to social insects, empirical studies for some species have shown that social learning occurs [26,28–31]. The fundamental idea of social learning is that an individual observes other individuals and changes their behaviour based on the others' presence or behaviour. This is a very broad notion. Individuals can be directly influenced by the observed behaviour of other individuals (learning), or they can be influenced by environmental social cues, such as pheromones or simply the presence of others [32–34]. Which behaviours are governed by which type of social influence is generally not well understood. In this study, we are only concerned with the dynamics of behaviour learned through imitation, which is already complex in itself and has not yet been widely investigated with mathematical models for social insects. Combining this with other social cues, for example pheromones, is a matter for future extensions of this framework. Thus, in our context, we apply the specific meaning that an individual copies a behaviour that is observed in others. There is indisputable empirical evidence that this happens [26]. Direct interaction or observation is necessary for this to occur. Given the possible complexity of social information exchange, we do not make any assumptions about its underlying mechanisms. We simply posit that individuals are more likely to imitate the behaviours of those who are successful. Independently of this, each individual may explore new behavioural variations with some probability.

We have previously established that certain empirically observed characteristics of colony behaviour, including task specialisation and the emergence of inactive subgroups can arise as a by-product of social learning mechanisms [35]. However, despite the well-established existence of social interactions in colonies, it is uncertain whether these behavioural phenomena can be attributed to mechanisms of social learning with certainty. This is because the exact scope and extent of social learning in insect colonies are not yet well understood. In this study, we investigate whether the same inactivity can also emerge if we only assume individual learning mechanisms.

We juxtapose pure *individual learning* and pure *social learning*, as two distinct methods of information processing by individuals at opposite ends of the spectrum of learning methods. Being based on very different types of information, these learning modes require very different cognitive and sensory capacities. Through a thorough examination of these two extremes, the study aims to gain an understanding of the effect that varying learning assumptions may have on the dynamics of the system.

Our study is motivated by the empirical evidence indicating the presence of both social and individual learning mechanisms in social insects. Bumble bees are a common model system that demonstrate both types of learning. An instance of this is the flower selection behavior in bumblebees, which can be a result of both individual learning and behavior copying [28] and bumblebees exhibit the ability to learn when to use each type of information [36]. We thus need to understand the differences and similarities of these types of learning mechanisms, including their comparative advantages and disadvantages for colony fitness.

To analyse the development of behaviour in a population under the social learning assumption, we employ adaptive dynamics, an analytic approach that originated in Evolutionary Game Theory (EGT) [37–39]. Agents follow basic rules to adjust their behaviour in response to an environmental signal, typically referred to as payoff [40]. Adaptive dynamics describes how a group responds to changes by taking into account the actions and interactions of individuals [41]. Evolutionary game theory was initially developed to model changes across evolutionary timescales, where the payoff represents fitness. However, this conceptual framework is not restricted to this timescale and can also be used to model faster processes that involve changes on colony lifetime scales, where payoffs are interpreted as feedback signals instead of fitness [35,42,43]. Our study is explicitly only concerned with these colony lifetime timescales. Interpreted in this way, adaptive dynamics captures how agents modify their task selection by taking into account task performance experience and environmental factors when working together on multiple tasks.

On the opposite end of the spectrum lies individual learning. It is commonly agreed that individual learning plays a crucial role [3,44,45] for social insects. It enables individuals to adjust to changing environments and improve their task performance over time. This is vital for the colony's survival, as it allows individuals to adapt to new challenges and make better decisions about how to allocate resources and solve problems. Individuals can adapt their strategies by utilising previously acquired information in their current context [30,46]. The arguably best-established model of task selection in social insects, the reinforcement response threshold model, is centrally based on this notion [47–50].

In this study, we employ a particularly well-studied form of Reinforcement Learning (RL) [51,52] where agents update their action probabilities using the cross rule of RL [53]. Cross-learning is a relatively simple type of reinforcement learning that is based on individual behaviour and fully aligns with the assumptions of the established adaptive threshold reinforcement model.

We compare the behavioural dynamics under these two learning assumptions for different types of environments. Our central aim is to investigate whether specific types of dynamics can be attributed to a specific learning mechanism, i.e., if they only emerge from social learning but not from individual learning or vice versa. We implement both processes in agent-based models to compare the outcomes. We back up the simulation studies with analytic results derived from adaptive dynamics.

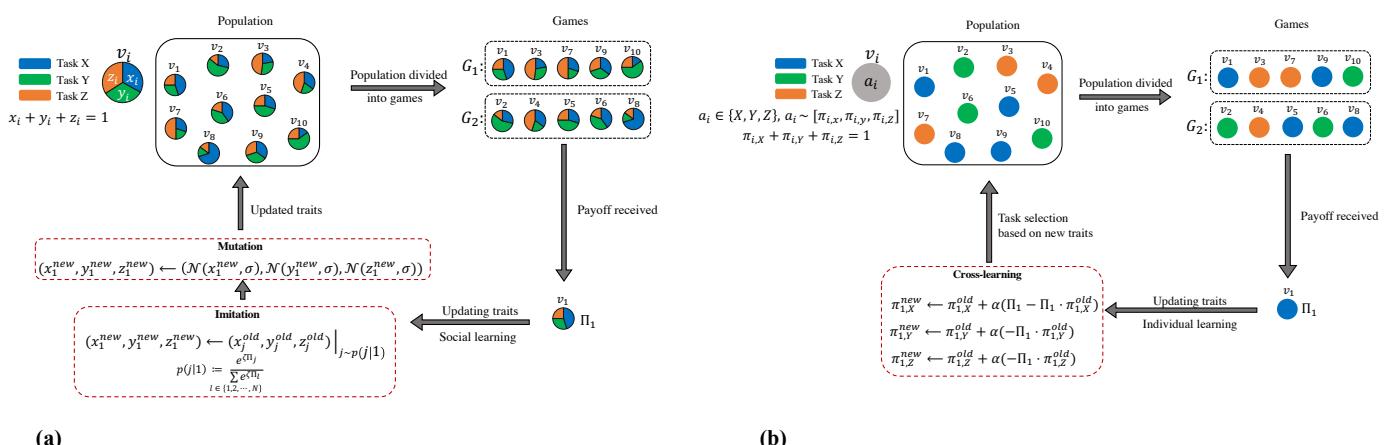
We are specifically interested in an effect previously referred to as *laziness* or inactivity in the population. This refers to the fact that in numerous efficient colonies, a significant portion of the workforce is comprised of inactive workers. This is a frequent occurrence in social organisations, including social insects, animals, humans, etc., which has been observed empirically and explained through modelling studies [35,54–58].

Our results show that identical behavioural dynamics, including the emergence of inactive workers, are observed *independently of the learning mode*. We conclude that this

inactivity can be a by-product of the collective learning process in a joint environment but is not conditioned by a particular type of learning.

## 2. Materials and Methods

We commence with a straightforward division of labour problem that only involves the selection of three prototypical tasks, which are labelled X, Y, and Z. To briefly summarise the core of social and individual learning frameworks, we make the assumption that there is a population of agents with a size of  $N$ , where each agent is entirely characterised by a set of trait values. Each model operates in discrete time steps. In each step, agents engage in group interactions of a predetermined size of  $n$  (known as “ $n$ -player games” in game theory terminology), and the population consists of  $K$  distinct  $n$ -player games,  $G_1, G_2, \dots, G_K$ , where  $K = \frac{N}{n}$ . Agent  $i$  in game  $G_k$  obtains a payoff  $\Pi_{i,G_k}$  from the group interaction, which is typically influenced by both the trait values of agent  $i$  and those of the other agents participating in the game. Nonetheless, the mechanism for learning (updating rule) varies depending on the type of learning. With *social learning*, agents acquire knowledge from one another by imitating or adopting the traits of another agent. Note that this can be viewed as being influenced by recruitment and imitation. In the event that the recruitment effort is adjusted based on task performance experience, proficient agents are more likely to be imitated. On the other hand, in *individual learning*, instead of imitation, each individual exploits their own experience and reinforces the probability of engaging in a certain task when the individual engages in it successfully. Figure 1 illustrates the general schematic of the dynamic process of both mechanisms.



**Figure 1.** (a) A diagrammatic representation of various stages in the social learning paradigm from an EGT perspective using a sample population with size  $N = 10$ . Each circle symbolises an individual in the population, and its segmentation into three colors denotes the distribution of involvement in the three tasks or response traits:  $x$ ,  $y$ , and  $z$ . The various steps of the process are demonstrated for an example agent,  $v_1$ , in the population. (b) Schematic diagram of different steps in the individual learning paradigm given three possible task choices. Each circle represents an individual and each colour represents the selected task by that individual. The different steps of the process are illustrated for a sample agent,  $v_1$ , in the population.

### 2.1. Social Learning Setting

To study the transition of behaviour in a population under social learning assumptions, we use adaptive dynamics as a framework of evolutionary game theory. Formally, agent  $i$  is characterised by a triple  $(x_i, y_i, z_i)$ , where the trait values  $x$ ,  $y$ , and  $z$  can be interpreted as the average fraction of effort invested into the first, second, and third task, respectively. As  $\forall i : x_i + y_i + z_i = 1$ , we can model a population of  $N$  workers as a two-dimensional vector of trait values  $(x_j, y_j)_{j=1, \dots, N}$ . As we are predominantly interested in the emergence of inactivity, we model inactivity as a third “pseudo-task” that does not generate benefit and

has no cost (see [35]). We thus have to have two “normal” tasks ( $X$  and  $Y$ ) with collective benefits and inactivity as a third “pseudo” task ( $Z$ ).

In numerous social organisations, such as social insects, the benefits arising from task completion are shared and are contingent on the cumulative effort invested, rather than solely on individual effort. An essential characteristic that we intend to investigate is task combinations in which a suitable number of workers need to perform multiple tasks to ensure the smooth functioning of the colony. Examples include brood care and thermoregulation. We model this with a multiplicative coupling of benefit  $B_X$  of Task  $X$  and  $B_Y$  of Task  $Y$  as follows:

$$B(\mathcal{X}_k, \mathcal{Y}_k) = \frac{1}{n} B_X\left(\sum_{x \in \mathcal{X}_k} x\right) \cdot B_Y\left(\sum_{y \in \mathcal{Y}_k} y\right). \quad (1)$$

where  $\mathcal{X}_k, \mathcal{Y}_k$  are the collective engagement levels of all individuals in  $G_k$ . The direct and immediate cost of executing a task, on the other hand, is borne by the agent performing the task and depends on the individual effort invested. Here, the third task ( $Z$ ; the level of inactivity) is assumed to cause no cost for the individuals. Hence, costs for multiple tasks are additive as below:

$$C(x_j, y_j) = C_X(x_j) + C_Y(y_j). \quad (2)$$

The payoff for individual  $j$  participating in game  $G_k$  is given as the difference between the benefit obtained and the cost incurred. In game  $G_k$ , individual  $j$  thus receives a payoff  $\Pi_{j,G_k}$  as:

$$\Pi_{j,G_k} = B(\mathcal{X}_k, \mathcal{Y}_k) - C(x_j, y_j). \quad (3)$$

The shape of the cost and benefit functions reflect the properties of the tasks and the environment. Details are given in Appendix A. (We follow [35]:  $X$  is a task with a concave benefit shape and marginally decreasing cost such as thermoregulation tasks in an ant colony;  $Y$  is a task with sigmoidal (thresholding) benefit shape and marginally decreasing costs such as brood care or defence tasks; and  $Z$  indicates inactivity (forgone effort), which produces no benefit and bears no cost.) Appendix B also explains the theoretical analysis and updating rules of adaptive dynamics. More details on the update rule associated with the Cross-learning are in Appendix C.

## 2.2. Individual Learning Setting

We analyse reinforcement learning, which is considered one of the simplest and most widely studied forms of individual or experience-based learning models. Reinforcement learning is a process in which an agent modifies its internal mixed strategy, which represents its behavioural disposition, modelled by a set of probability distributions determining how individual actions are selected. If a task execution results in a high payoff in the past, its future probability increases, reinforcing the behaviour associated with the action. (This is akin to lowering the threshold in the reinforced threshold model). Reinforcement protocols have substantial empirical support and have been widely used to model a range of complex behaviours in social and biological systems [59]. We study RL in a population game as an appropriate representation of collective behaviour modification in a colony.

We employ a particular form of RL where agents update their action probabilities using the cross rule of RL [53]. Cross-learning is a straightforward form of individual-based reinforcement learning that aligns with the widely accepted threshold reinforcement model for the division of labour in social insects.

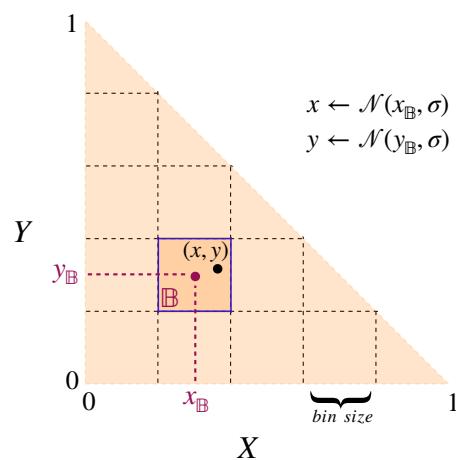
For the cross-learning framework, the most intuitive choice would be to characterise each worker by the probability with which she engages in a particular task. Agent  $i$  would thus be represented by a triple  $(\pi_{i,X}, \pi_{i,Y}, \pi_{i,Z})$  where  $\pi_{i,X}, \pi_{i,Y}$ , and  $\pi_{i,Z} \in \mathcal{R}$  are the probabilities of executing the first, second, and third tasks, respectively ( $\forall i : \pi_{i,X} + \pi_{i,Y} + \pi_{i,Z} = 1$ ).

Note that this would imply an important difference in modelling the social learning and the individual learning processes. Instead of dividing the invested effort between three tasks (as in the social learning paradigm), each individual fully engages in a single task with a probability given by its trait values. Thereby, given the conventional form of cross-learning, we can only account for discrete task engagement patterns for individuals. However, to account for the possibility of non-binary participation in tasks that involve continuous trait values, which are commonly used in social learning, we extend the conventional cross-learning algorithm. Instead of having a binary choice for selecting a task or not, we model the level of engagement by discretising the levels of engagement into bins. Each bin corresponds to a pair of ranges for both  $x$  and  $y$  traits, model the proportion of effort invested in the tasks, exactly as in the social learning paradigm, and each individual bin is assigned a probability of being selected. One might argue that the level of engagement (in social learning) could be interpreted as the long-term average of the task execution frequency. While this is a reasonable interpretation, this will only result in comparable payoffs if the expectation of the payoff of individual task engagements is identical to the payoff of the expected level of engagement. This is generally not the case for non-linear payoff functions. Here, we choose bins of size  $0.05 \times 0.05$  resulting in 210 different pairs of value ranges for  $x$  and  $y$  traits.

Figure 2 illustrates the binarising process in the proposed modified version of the cross-learning algorithm.

Then, for the purpose of comparison, we can define the multiplicative coupling of benefit  $B_X$  of Task X and  $B_Y$  of Task Y and immediate costs of  $C_X$  and  $C_Y$  similar to the previous setting in Section 2.1.

More details on the update rule associated with the adaptive dynamics can be found in Appendix B.



**Figure 2.** Illustrating the process of binarising the space of possible  $(x, y)$  choices in the modified cross-learning algorithm. At each step of updating the trait value pair,  $(x, y)$ , a single bin  $\mathbb{B}$  is first selected. The new  $(x, y)$  pair is then calculated, applying a random Gaussian distribution around the center point of the bin, which is shown as  $(x_{\mathbb{B}}, y_{\mathbb{B}})$ .

### 3. Results

We compared the behavioural trajectories of the models discussed above, dependent on parameters  $b_1$ ,  $b_2$ ,  $w$ , and  $\beta$ , which are embedded in the benefit and cost functions and reflect the properties of the environment (see Table 1). We implemented the models as individual-based, discrete-time simulations starting from a monomorphic population. At each time step, the population was randomly divided into  $K$  sets of  $n$  individuals each, where  $n$  is a fixed group size. Each individual received a payoff determined by their trait values and the composition of the group they are part of. In the case of social learning, the individuals were then recruited to successful behaviours (technically, individuals imitate the trait values of others with a probability that is determined by the recruiter's performance

compared to the average performance of the entire population). Each trait value could also undergo a slight change that could be considered an autonomous exploration of behaviour through variation, similar to a mutation. In individual learning, due to the update rules related to cross-learning, each agent modifies their trait values at each time step by reinforcing the probability of the action executed according to the task-related reward experienced. Full details of each method are given in Algorithms 1 and 2.

**Table 1.** Model parameters.

Notation	Definition	Interpretation	Values
$b_1$	linear benefit coefficient for task X (e.g., thermoregulation)	efficiency of regulation	{16, 20, 28}
$b_2$	quadratic benefit coefficient for task X	shape of benefit for task X (homeostatic vs. maximising)	{−4, −6}
$\beta$	slope of benefit for task Y (e.g., brood care)	larger values indicate higher efficiency per unit of work	3
$w$	inflection point of benefit for task Y	$1/w$ indicates minimum amount of work required	{0.2, 0.3, 0.5}
$n$	group size for individual interactions		5
$\zeta$	selection intensity in social learning		2
$\mu$	mutation rate in social learning	probability of behaviour exploration	0.01
$\sigma$	mutation size in social learning	amount of behaviour variation	0.005
$\alpha$	learning rate (step size) in individual learning	determines to what extent newly acquired information overrides old information	0.01

Figure 3 shows the simulation results of both models for different sets of environmental parameters. (The source code of the simulations is available at this [GitHub link](#)).

The different behaviour variations were classified into three groups:

**Fully specialised:** Regardless of the boundary conditions, the entire population uniformly shifts towards full engagement in a single task (task Z or inactivity), resulting in inviability.

**Branching:** After initial movement toward a shared level of engagement, the population splits into two (or more) co-existing traits. These sub-populations show different levels of engagement in the three tasks.

**Uniform behaviour; fully generalised:** In this case, all individuals move toward a shared level of engagement in each of the three tasks (i.e., a shared set of all trait values with a certain level of inactivity). From the EGT perspective, this represents an Evolutionary Stable Strategy (ESS).

The simulation results in Figure 3 illustrate that both learning paradigms resulted in the same behaviour in all behavioural environments. Given a monomorphic initial population, all individuals first move towards a fixed point (red dot in the streamline plots) starting from an initial set of trait values (green dot in the streamline plots). This fixed point was predicted analytically using adaptive dynamics as shown in the streamline plots of Figure 3. In certain environments, a branching behaviour occurs after reaching the fixed point. This split is also predicted by adaptive dynamics for the case of social learning. The simulations show that this is not unique to social learning but that both individual and social learning dynamics exhibit the same split (see Left and Right simulation results in Figure 3b for individual learning and social learning results respectively). More intriguingly, the results depict that both learning mechanisms can simulate the emergence of inactivity (i.e., non-zero engagement in task Z at the steady state) in certain parameter ranges. Thus, the models suggest that under certain environmental conditions, inactivity can arise simply as a by-product of the collective adjustment process in a joint environment without being restricted to a specific form of learning.

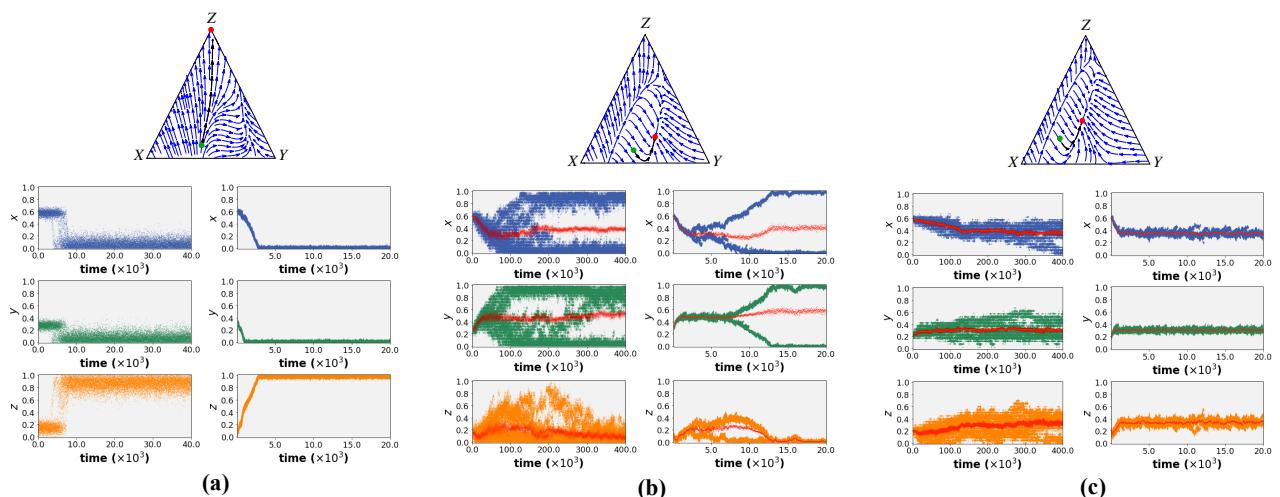
**Algorithm 1** SOCIAL LEARNING: At each generation  $t$ , each individual  $j$  updates its strategy  $(x_j^t, y_j^t)$  following an imitation phase and a mutation with probability  $\mu$

---

**Require:** A population of size  $N$  with a strategy profile  $p^0 = \{(x_1^0, y_1^0), (x_2^0, y_2^0), \dots, (x_N^0, y_N^0)\}$  at generation  $t = 0$ ;  $(x_j^0, y_j^0) = (x_0, y_0) \in ([0, 1], [0, 1])$ ; selection intensity,  $\zeta$ ; mutation rate,  $\mu$ ; standard deviation for Gaussian mutations,  $\sigma$ .

- 1:  $P \leftarrow []$
- 2: **for**  $t = 1 : T$  **do**
- 3:      $P.append(p^{t-1})$
- 4:     generate  $K = \frac{N}{n}$  random games
- 5:     **for**  $j = 1 : N$  **do**
- 6:          $\Pi_{j,G_k} \leftarrow B(\mathcal{X}_k^{t-1}, \mathcal{Y}_k^{t-1}) - C(x_j^{t-1}, y_j^{t-1})$
- 7:         where  $j \in G_k; k \in \{1, \dots, K\}$
- 8:     **end for**
- 9:     **for**  $j = 1 : N$  **do**
- 10:         imitates  $l$  with probability  $\frac{e^{\zeta \Pi_l}}{\sum_{m=1}^N e^{\zeta \Pi_m}}$
- 11:          $x_j^t \leftarrow x_l^{t-1}$
- 12:          $y_j^t \leftarrow y_l^{t-1}$
- 13:     **end for**
- 14:     **for**  $j = 1 : N$  **do**
- 15:         **if**  $\text{random}() < \mu$  **then**
- 16:              $x_j^t \leftarrow \max(0, \min(\mathcal{N}(x_j^t, \sigma), 1))$
- 17:              $y_j^t \leftarrow \max(0, \min(\mathcal{N}(y_j^t, \sigma), 1))$
- 18:         **end if**
- 19:     **end for**
- 20:      $p^t \leftarrow \{(x_1^t, y_1^t), (x_2^t, y_2^t), \dots, (x_N^t, y_N^t)\}$
- 21: **end for**
- 22: **Return**  $P = \{p^0, p^1, \dots, p^{T-1}\}$

---



**Figure 3.** Streamline plots, as well as **Left:** individual learning, and **Right:** social learning simulation results for different environmental settings. The streamline depicts the development of a monomorphic population from the initial point (in green) to the fixed point (in red). **(a) Fully specialised:**  $b_1 = 16, b_2 = -6, \omega = 0.2$ . **(b) Branching:**  $b_1 = 20, b_2 = -4, \omega = 0.3$ . **(c) Uniform behaviour; fully generalised:**  $b_1 = 28, b_2 = -6, \omega = 0.5$ . The average population trait values are plotted in red which closely match between individual learning (cross-learning simulations) and social learning (adaptive dynamics simulations).

---

**Algorithm 2** MODIFIED CROSS-LEARNING: At each generation  $t$ , each individual  $j$  updates its strategy  $(x_j^t, y_j^t)$  and probability distribution  $(\pi_{j,1}^t, \dots, \pi_{j,m}^t)$  over  $m$  possible action bins  $\mathbb{B}_1, \dots, \mathbb{B}_m$  with learning rate of  $\alpha$

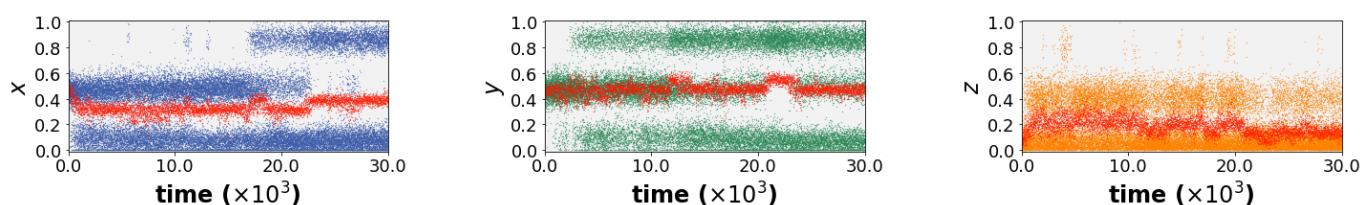
---

**Require:** A population of size  $N$  with a strategy profile  $p^0 = \{(x_1^0, y_1^0), (x_2^0, y_2^0), \dots, (x_N^0, y_N^0)\}$  and probability distribution profile  $\Pi^0 = \{(\pi_{1,1}^0, \dots, \pi_{1,m}^0), \dots, (\pi_{N,1}^0, \dots, \pi_{N,m}^0)\}$  at generation  $t = 0$ ;  $(x_j^0, y_j^0) = (x_0, y_0) \in \mathbb{B}_l$ ;  $\pi_{j,l}^0 = 1, \pi_{j,a}^0 = 0 \quad \forall a \neq l$ ; learning rate,  $\alpha$ .

- 1:  $P \leftarrow []$
- 2: **for**  $t = 1 : T$  **do**
- 3:      $P.append(p^{t-1})$
- 4:     generate  $K = \frac{N}{n}$  random games
- 5:     **for**  $j = 1 : N$  **do**
- 6:         agent  $j$  chooses an action bin
- 7:          $a_j^{t-1} \in \{\mathbb{B}_1, \dots, \mathbb{B}_m\}$
- 8:          $(x_{a_j^{t-1}}, y_{a_j^{t-1}}) \leftarrow$  center point of  $a_j^{t-1}$
- 9:          $x_j^{t-1} \leftarrow \max(0, \min(\mathcal{N}(x_{a_j^{t-1}}, \sigma), 1))$
- 10:          $y_j^{t-1} \leftarrow \max(0, \min(\mathcal{N}(y_{a_j^{t-1}}, \sigma), 1))$
- 11:          $I_j = l \quad s.t. \quad (x_j^{t-1}, y_j^{t-1}) \in \mathbb{B}_l, \quad l \in \{1, \dots, m\}$
- 12:     **end for**
- 13:     **for**  $j = 1 : N$  **do**
- 14:          $\Pi_{j,G_k} \leftarrow B(\mathcal{X}_k^{t-1}, \mathcal{Y}_k^{t-1}) - C(x_j^{t-1}, y_j^{t-1})$
- 15:         where  $j \in G_k; \quad k \in \{1, \dots, K\}$
- 16:     **end for**
- 17:     **for**  $j = 1 : N$  **do**
- 18:          $\pi_{j,l}^t \leftarrow \pi_{j,l}^{t-1} + \begin{cases} \alpha(\Pi_{j,G_k} - \Pi_{j,G_k} \cdot \pi_{j,l}^{t-1}) & l = I_j \\ \alpha(-\Pi_{j,G_k} \cdot \pi_{j,l}^{t-1}) & o.w \end{cases}$
- 19:     **end for**
- 20:      $p^t \leftarrow \{(x_1^t, y_1^t), (x_2^t, y_2^t), \dots, (x_N^t, y_N^t)\}$
- 21: **end for**
- 22: **Return**  $P = \{p^0, p^1, \dots, p^{T-1}\}$

---

Finally, we repeat our analysis in the same environmental setting as the branching region in Figure 3b using the modified cross-learning framework but with a larger bin size of 0.2. This divides the entire space of possible  $x$  and  $y$  paired value ranges to six possible bins. We coarse-grain the model in order to address the concern that a fine-grained bin model is surely not biologically plausible. What is plausible, though, is that the individual may have some concept of executing a task “always”, “never”, “frequently”, or “infrequently”. This is adequately captured in the coarse-grained bin model. As expected, the results in Figure 4 show qualitatively the same behaviour as in Figure 3 but with noise added. These findings confirm the fact that the findings do not change under a cognitively plausible model of trait values.



**Figure 4.** Individual learning simulation results with bin sizes equal to 0.2 in the **branching** region:  $b_1 = 20$ ,  $b_2 = -4$ ,  $\omega = 0.3$ . This binarisation results in 6 different pairs of value ranges for  $x$  and  $y$  traits; the settings mimic a case of 6 different tasks in an unmodified cross-learning framework.

#### 4. Discussion

Division of labour is essential for the survival and ecological success of social organisations. By dividing tasks among individuals, a social organisation can ensure that the most skilled or efficient individuals are performing specific tasks, which can lead to increased productivity and overall success. Dividing labour can also promote specialisation in the population, as individuals are able to focus on specific tasks and develop expertise in those areas. Furthermore, it also allows for flexibility and the ability to respond quickly to changes in the environment and internal requirements.

Social insect colonies are examples of the most ecologically successful life forms, and an efficient division of labour is a critical aspect of their success. There has been a significant amount of research on the division of labour in social insects [1]. However, much of this research has focused on the impact of internal factors such as genetics [60], morphology [61], and hormones [62]. In comparison, there has been relatively less focus on the impact of the environment on task choices at the individual level and the underlying mechanisms of social interactions and their role in regulating the division of labour.

We studied two of the most widely used methods for modelling the mechanisms of the division of labour in social organisations: social learning and individual learning. Very few previous studies have focused on comparing the similarities and differences in the outcomes resulting from the different update rules used. A comprehensive comparison of the two frameworks in various environmental settings is crucial in understanding the advantages and limitations of each assumption. It will help in the better understanding of the underlying mechanisms of the division of labour, in general, and more specific phenomena such as the emergence of inactivity as observed in empirical data, in particular.

In this study, we have attempted to gain a deeper understanding of the implications of these two different learning paradigms for a specific behavioural phenomenon observed in social colonies: the emergence of inactive subgroups that do not participate in the collective action that sustains the colony.

Previous studies have posited that this particular phenomenon, an instance of branching behaviour, is a by-product of the learning mechanism [35]. However, it was unclear whether these aspects of dynamics were indeed influenced by the presence of social interactions and the learning mechanism itself. By comparing and contrasting the results of both individual and social learning paradigms across different environmental conditions, we found equivalent behavioural outcomes in both cases. Specifically, we demonstrated that an individual, experience-based learning approach can also lead to inactivity in the population. This supports the hypothesis that regardless of the dominant learning mechanism in the colony, this aspect of colony life can arise as an artefact of the collective behaviour modification in a joint environment but is not necessarily restricted to a specific learning mechanism.

Using mathematical intuition, this is not entirely surprising. There are deep correspondences between cross-learning and social learning. The seminal contribution by Borgers [63] first established that cross-learning and replicator dynamics (Appendix D), a formal model of learning by imitation, exhibit similar dynamics. However, the important restriction of this result is that it only applies to the learning of a single individual and that it can only be proven *in expectation* and in the continuous-time limit. The setting of a social insect

colony, however, is necessarily population-based learning. The term population-learning is adapted from the reinforcement literature and can refer to any interacting collective, rather than to the specific meaning of population in biology. Börgers and Sarin's finding was later extended by Lakhar and Seymour to show that population-based cross-learning evolves according to a specific form of the replicator equation under certain conditions, the so-called replicator continuity equation [64]. This equation is a partial differential equation that describes the changes in the population state over time. This was a very important finding, but it is restricted to replicator dynamics, which can only capture discrete behavioural states.

Adaptive dynamics, which we have used here, can capture continuous behaviour parameters (such as an engagement level) and can analytically predict whether a population will split into subgroups. What we have demonstrated in this paper is that adaptive dynamics and population-based cross-learning exhibit qualitatively equivalent dynamics in the context of the study of inactivity.

The presence of the studied inactive subgroups is a common occurrence in collective behavior, observed in organisations such as active particles, insects, animals, and humans [55,65]. Social insect colonies, in particular, can have over half of their workers inactive at a given time [66], which is surprising given the low individual selfishness levels in these colonies [13]. Despite our hypothesis that inactivity can arise as a by-product of the task allocation process and independent of the learning mode, in certain situations, this inactivity has been proposed to have a functional purpose. The main hypothesis for the functional role of inactive workers is that they serve as a reserve workforce that can be mobilised quickly when there is a sudden loss of workers or unexpectedly high task demands, thereby increasing colony flexibility and resilience [67,68]. Nevertheless, the benefits derived from having a reserve workforce of inactive individuals have never been quantified, either empirically or otherwise, leading many empirical research to still question this hypothesis [69]. Examples of other explanations include sleep or rest time of individuals [70,71], or delays occurring during task switching and the time the workers require to asses the collected information about task demands without engaging in any work [72]. However, the variation among individuals in social insect colonies in terms of their amount of inactivity cannot be fully explained by the need for resting periods alone [13] and although challenging, the purposeful activity of searching for a task, or "patrolling" should be distinguished from aimless and inactive wandering [73]. This suggests the necessity for additional research on the subject, both through empirical studies and theoretical or modeling work, as proposed in this study.

Perhaps the main limitation of our study is that we have solely looked at social learning and individual learning in isolation. Yet, it is likely that, in many circumstances, both may occur simultaneously and even intermingle, possibly for individual task selection or even in a context-dependent manner, as shown in previous research [36]. While we have focused on analyzing isolated forms of each learning paradigm, investigating such mixed modes is a task for future studies. Our paper's primary objective was to demonstrate that the overall behavior dynamics in the population are similar under both learning paradigms. Therefore, it is probable that a combination of learning mechanisms would also manifest similar dynamics. It should be relatively straightforward to confirm this with simulations, however a mathematical framework that captures both modes simultaneously is unclear.

To conclude, we believe that our approach, beyond the study of this particular phenomenon of collective behaviour (i.e., inactivity), hints at new pathways for studying collective behaviour in animal groups from a generalised perspective without having to assume (or know) a restrictive model of learning. To explore this possibility further, the exact conditions under which these equivalences hold will have to be established formally and we hope to do so in future work.

**Author Contributions:** Conceptualisation, M.K. and B.M.; methodology, M.K. and B.M.; software, M.K.; validation, M.K., B.M. and J.G.; formal analysis, M.K.; investigation, M.K.; writing—original draft preparation, M.K.; writing—review and editing, M.K., B.M. and J.G.; visualisation, M.K.; supervision, B.M. and J.G.; project administration, B.M.; funding acquisition, B.M. and J.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Australian Research Council; ARC Discovery (DP180100154, DP200100036).

**Data Availability Statement:** The scripts and simulations supporting the reported results can be found in [https://github.com/Moein-Khajehnejad/Social\\_learning-vs.-Individual\\_learning](https://github.com/Moein-Khajehnejad/Social_learning-vs.-Individual_learning).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

EGT	Evolutionary Game Theory
ESS	Evolutionary Stable Strategy
RL	Reinforcement Learning

## Appendix A. Benefit and Cost Functions

The state of a population with size  $N$  is fully determined by a 2-dimensional vector  $(x_j, y_j)_{j=1,\dots,N}$ , where  $x_j$  and  $y_j$  are the engagement levels in the two tasks  $X$  and  $Y$ , respectively. Trivially,  $z_j = 1 - x_j - y_j$  is the level of individual  $j$  in task  $Z$ . The interactions among individuals are limited to groups of size  $n$  so that we have  $K = \frac{N}{n}$  groups or games. Each individual's payoff depends not only on their own trait values but also on the trait values of all other players in the same game. Particularly, a game  $G_k$ ,  $k \in \{1, \dots, K\}$  induces trait vectors  $\mathcal{X}_k = \{x_i\}_{i \in G_k}$  and  $\mathcal{Y}_k = \{y_i\}_{i \in G_k}$  (engagement levels of all individuals in the game in tasks  $X$  and  $Y$ , respectively).

For individual  $j$  in the game  $G_k$ , the payoff is:

$$\Pi_{j,G_k} = B(\mathcal{X}_k, \mathcal{Y}_k) - C(x_j, y_j) \quad (\text{A1})$$

where  $B(\mathcal{X}_k, \mathcal{Y}_k)$  is the total benefit in game  $G_k$  shared evenly among  $n$  individuals and  $C(x_j, y_j)$  is the individual costs for the worker  $j$ .

Let  $B_X(\cdot), B_Y(\cdot)$ , and  $C_X(\cdot), C_Y(\cdot)$  be the benefit and cost functions associated with tasks  $X$  and  $Y$ , respectively. We assume both tasks are necessary for the colony's fitness and represent this with a multiplicative function for the total benefit. Costs, on the other hand, are naturally additive.

$$B(\mathcal{X}_k, \mathcal{Y}_k) = \frac{1}{n} B_X\left(\sum_{i \in G_k} x_i\right) \cdot B_Y\left(\sum_{i \in G_k} y_i\right) \quad (\text{A2})$$

$$C(x_j, y_j) = C_X(x_j) + C_Y(y_j) \quad (\text{A3})$$

The benefit of a homeostatic task is adequately modelled with a concave quadratic function, reflecting the fact that a certain intermediate level of investment of effort is optimal. For example, allocating too little or too much effort can lead to inefficient regulation.

$$B_X\left(\sum_{i \in G_k} x_i\right) = \text{Max}\left[b_2 \left(\sum_{i \in G_k} x_i\right)^2 + b_1 \left(\sum_{i \in G_k} x_i\right), 0\right] \quad (\text{A4})$$

$$C_X(x_j) = -x_j^2 + 2x_j \quad (\text{A5})$$

The benefit function is lower-bounded to ensure non-negative benefit values. We have a marginally decreasing cost function for the regulation task capturing efficiency improvement through task experience ( $\forall j : C'_X(x_j) > 0$  and  $\forall j : C''_X(x_j) < 0$  ).

Parameters  $b_1$  and  $b_2$  reflect environmental influences. These parameters also determine the shape of  $B_X$  and can be adjusted to a homeostatic task or a maximising task with purely increasing benefits.

$Y$  is a thresholding task, e.g., brood care. For a given level of brood, a minimum level of labour is required. However, exceeding this level does not generate much additional benefit. The parameters  $\beta$  and  $w$  control the slope of the function and the value of the threshold, respectively. We assume  $C_Y$  to be a linear increasing function.

$$B_Y\left(\sum_{i \in G_k} y_i\right) = \frac{1}{1 + \left(\frac{1-w(\sum y_i)}{w(\sum y_i)}\right)^\beta} \quad (\text{A6})$$

$$C_Y(y_j) = -y_j^2 + 2y_j \quad (\text{A7})$$

## Appendix B. Update Rule of Adaptive Dynamics

In the environment consisting of three tasks  $X$ ,  $Y$ , and  $Z$ , two traits  $x$  and  $y$  change jointly and  $z = 1 - x - y$ . Assume an  $n$ -player continuous game with players  $((x_1, y_1), \dots, (x_n, y_n))$ . We can then obtain the invasion fitness (growth rate) of a rare mutant with strategy  $(x', y')$  among the resident population with strategy  $(x, y)$  ( $x, y, x', y' \in [0, 1]$ ) as follows:

$$\begin{aligned} S_{x,y}(x', y') &= \Pi((x', y'); [(x, y)]_{n-1}) - \Pi((x, y); [(x, y)]_{n-1}) \\ \text{s.t. } [(x, y)]_n &= (\underbrace{(x, y), \dots, (x, y)}_n), \end{aligned}$$

where

$$\Pi((x', y'); [(x, y)]_{n-1}) = \frac{1}{n} B_X(x' + (n-1)x) \cdot B_Y(y' + (n-1)y) - C_X(x') - C_Y(y')$$

is the expected payoff of the rare mutant  $(x', y')$  and

$$\Pi((x, y); [(x, y)]_{n-1}) = \frac{1}{n} B_X(nx) \cdot B_Y(ny) - C_X(x) - C_Y(y)$$

is the expected payoff of the monomorphic population with strategy  $(x, y)$ . Traits  $x$  and  $y$  will then be updated given the following selection gradients:

$$\begin{aligned} \mathcal{D}_x(x, y) &= \left. \frac{\partial S_{x,y}(x', y')}{\partial x'} \right|_{\substack{x'=x \\ y'=y}} \\ &= \frac{1}{n} B'_X(nx) \cdot B_Y(ny) - C'_X(x) \end{aligned}$$

$$\begin{aligned} \mathcal{D}_y(x, y) &= \left. \frac{\partial S_{x,y}(x', y')}{\partial y'} \right|_{\substack{x'=x \\ y'=y}} \\ &= \frac{1}{n} B_X(nx) \cdot B'_Y(ny) - C'_Y(y). \end{aligned}$$

Singular coalitions are then found by the concurrent solutions of  $\mathcal{D}_x(x^*, y^*) = 0$  and  $\mathcal{D}_y(x^*, y^*) = 0$ .

The stability of a singular coalition  $(x^*, y^*)$  can now be determined by performing a linear stability analysis around  $(x^*, y^*)$  and examining the sign of the eigenvalues of the Jacobian of the selection gradients ( $J$ ).

$$\mathbf{J} = \begin{pmatrix} \frac{\partial D_x(x,y)}{\partial x} \Big|_{\substack{x=x^* \\ y=y^*}} & \frac{\partial D_x(x,y)}{\partial y} \Big|_{\substack{x=x^* \\ y=y^*}} \\ \frac{\partial D_y(x,y)}{\partial x} \Big|_{\substack{x=x^* \\ y=y^*}} & \frac{\partial D_y(x,y)}{\partial y} \Big|_{\substack{x=x^* \\ y=y^*}} \end{pmatrix}$$

Let  $\lambda_1$  and  $\lambda_2$  be the eigenvalues of  $\mathbf{J}$ . The singular coalition  $(x^*, y^*)$  is an attractor of the evolutionary dynamics (convergent stable) if both  $\lambda_1$  and  $\lambda_2$  have negative real parts ( $\Re\{\lambda_1\}, \Re\{\lambda_2\} < 0$ ) and repelling (unstable) if at least one eigenvalue has positive real part ( $\Re\{\lambda_1\} \cdot \Re\{\lambda_2\} > 0$ ).

The outcome of the singular coalition and its susceptibility to invasion by rare proximate mutants is determined by the eigenvalues of the Hessian matrix ( $\mathbf{H}$ ) of the invasion fitness:

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 S_{x^*,y^*}(x',y')}{\partial x'^2} \Big|_{\substack{x'=x^* \\ y'=y^*}} & \frac{\partial^2 S_{x^*,y^*}(x',y')}{\partial x' \partial y'} \Big|_{\substack{x'=x^* \\ y'=y^*}} \\ \frac{\partial^2 S_{x^*,y^*}(x',y')}{\partial y' \partial x'} \Big|_{\substack{x'=x^* \\ y'=y^*}} & \frac{\partial^2 S_{x^*,y^*}(x',y')}{\partial y'^2} \Big|_{\substack{x'=x^* \\ y'=y^*}} \end{pmatrix}$$

Since  $\mathbf{H}$  is symmetric, its eigenvalues  $\mu_1$  and  $\mu_2$  are real. Thus, if both  $\mu_1$  and  $\mu_2$  are negative, the population at  $(x^*, y^*)$  is uninvadable by nearby mutants and they reside at the maximum of the fitness landscape. If both  $\mu_1$  and  $\mu_2$  are positive,  $(x^*, y^*)$  is invadable by all rare nearby mutants. Lastly, if the product of the eigenvalues ( $\mu_1 \cdot \mu_2$ ) is negative, the singular coalition can be invaded by some mutants but not others, and the point  $(x^*, y^*)$  is a saddle point.

### Appendix C. Update Rule of Cross-Learning

We utilise the cross rule of learning [53], a specific form of reinforcement learning. This rule dictates that a decision maker will raise the probability of their chosen action in the previous round based on the payoff received, while simultaneously decreasing the probability of alternative actions in proportion. Previous works have analysed the cross rule and characterised it as the prototype of a class of RL schemes [63]. As shown in [63], the cross rule has a noteworthy implication; it states that the expected change in the probability of an action is equivalent to the replicator dynamic from evolutionary game theory.

In the original form of cross-learning (considering a binary choice for selecting/not selecting each present of the present tasks), we assume each individual  $i$  in the population is a cross-learner. Each individual then updates their action probabilities  $(\pi_{i,X}, \pi_{i,Y}, \pi_{i,Z})$  based on the payoff received after taking action  $a$  as follows

$$\pi_{i,k} \leftarrow \pi_{i,k} + \begin{cases} \alpha(\Pi_i - \pi_{i,k} \cdot \Pi_i) & \text{if } k = a \\ \alpha(-\pi_{i,k} \cdot \Pi_i) & \text{otherwise} \end{cases} \quad (\text{A8})$$

where  $k \in \{X, Y, Z\}$ ,  $\alpha$  is the learning step size, and  $\Pi_i$  is the received payoff of individual  $i$ .

The same update rule then extends to our modified version of cross-learning with the number of possible choices extended to the number of bins in the two-dimensional space of the two trait values  $x$  and  $y$ . Let us assume  $\{\mathbb{B}_1, \dots, \mathbb{B}_m\}$  constitutes the set of all the possible action bins. At each step, the bin with the highest probability is chosen. Then, the new  $x'$  and  $y'$  trait values are selected from within that bin by applying a random Gaussian distribution around the center point of the bin. Formally, we have:

$$\begin{aligned}x' &= \max(0, \min(\mathcal{N}(x_{m_l}, \sigma), 1)), \\y' &= \max(0, \min(\mathcal{N}(y_{m_l}, \sigma), 1)) \\z' &= 1 - x' - y',\end{aligned}\tag{A9}$$

where  $\mathbb{B}_l$  is the selected bin with the highest probability and  $(x_{m_l}, y_{m_l})$  is the center point of  $\mathbb{B}_l$ .

#### Appendix D. Replicator Dynamics

In contrast to classical game theory, which assumes players have complete knowledge of the game and act rationally, evolutionary game theory incorporates concepts from biology, such as natural selection and mutation, and relaxes the rationality assumption to better reflect the dynamic nature of real-world interactions [41]. A key aspect of evolutionary game theory is the replicator dynamics, which explains how a population of individuals changes over time under evolutionary pressure. Each individual belongs to a specific type and interacts with others randomly. Their reproductive success is based on their fitness, which is determined by these interactions. According to replicator dynamics, if individuals of a certain type have a higher fitness than the average population, the population share of that type will increase. Conversely, if their fitness is lower than the average, their population share will decrease.

Assuming a two-action (binary choice) game, we can employ the replicator dynamics to evaluate the fate of the population. In this case, at any time point  $t$ , the population can be described by the state vector  $\mathbf{S}_t = (s_1(t), s_2(t), \dots, s_m(t))$ , with  $0 \leq s_i(t) \leq 1, \forall i$  and  $\sum_{i=1}^m s_i(t) = 1$ , representing the fractions of the population dependent to each of  $m$  types. Replicator dynamics describe how a large population of individuals changes behaviour over time. However, they can also be interpreted as the strategy of a single player. In this interpretation, the population share of each type represents the probability that the player selects the corresponding pure action. The replicator dynamics then describe how the player's strategy changes over time as they repeatedly play the game and update their policy.

Now, in our environmental setting, let  $x(t), y(t)$  represent the fraction of agents from the entire population that engages at time  $t$  in task X and Y, respectively. Then, the fraction of individuals doing Z is given by  $(1 - x(t) - y(t))$ . So, the state space can be shown as  $\mathbf{S}_t = (s_1(t) = x(t), s_2(t) = y(t), s_3(t) = 1 - x(t) - y(t))$ . Based on the changes in the expected average payoffs for these tasks, the replicator dynamics provide an estimate of how the fraction of workers in each task changes over time as players observe the performance of others and attempt to improve their own relative performance. So, here, the expected average payoff to an individual doing task X is given by:

$$\Pi_X(t) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1-i} \binom{n-1}{i; j} x(t)^i \cdot y(t)^j \cdot (1 - x(t) - y(t))^{n-1-i-j} \cdot \left[ \frac{1}{n} B_X(i+1) \cdot B_Y(j) - C_X \right],\tag{A10}$$

and for a player doing Y, it is given by:

$$\Pi_Y(t) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1-i} \binom{n-1}{i; j} x(t)^i \cdot y(t)^j \cdot (1 - x(t) - y(t))^{n-1-i-j} \cdot \left[ \frac{1}{n} B_X(i) \cdot B_Y(j+1) - C_Y \right],\tag{A11}$$

and finally, for a player doing Z, it is

$$\Pi_Z(t) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1-i} \binom{n-1}{i; j} x(t)^i \cdot y(t)^j \cdot (1 - x(t) - y(t))^{n-1-i-j} \cdot \left[ \frac{1}{n} B_X(i) \cdot B_Y(j) \right],\tag{A12}$$

$$\text{where } \binom{n-1}{i; j} = \frac{n!}{i! \cdot j! \cdot (n-1-i-j)!}.$$

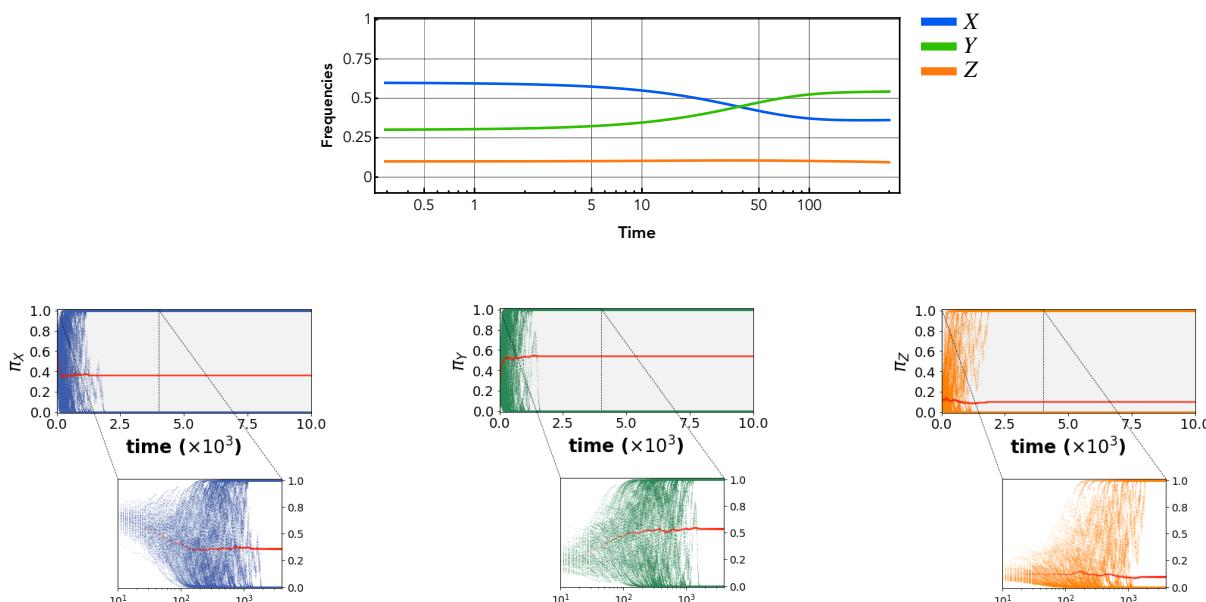
The expected payoff received by a randomly selected individual in the population is then given by

$$\bar{\Pi}(t) = x(t) \cdot \Pi_X(t) + y(t) \cdot \Pi_Y(t) + (1 - x(t) - y(t)) \cdot \Pi_Z(t). \quad (\text{A13})$$

The corresponding replicator dynamics for the fraction of  $x(t)$  and  $y(t)$  is given by two differential equations; each represents the fraction of the population at a specific task growing at a rate that is proportional to the difference between the payoff of that task to the average payoff of the entire population. Particularly, this is given by:

$$\begin{aligned} \dot{x}(t) &= x(t)(\Pi_X(t) - \bar{\Pi}(t)) \\ \dot{y}(t) &= y(t)(\Pi_Y(t) - \bar{\Pi}(t)). \end{aligned} \quad (\text{A14})$$

Figure A1 represents the simulation results of replicator dynamics in a sample environmental setting also explored in the main text. We compare these results with that of a conventional cross-learning, where only a binary choice is possible (i.e., individuals either choose a certain task or not). It is illustrated that the mixed population strategies are identical in both replicator dynamics and cross-learning as also implied in [63]. The same outcome is achieved for other environmental settings (non-branching) given we consider a binary choice and the conventional form of the cross-learning framework.



**Figure A1.** Simulation results of the system using replicator dynamics (**top**) and cross-learning (**bottom**) given:  $b_1 = 24$ ,  $b_2 = -6$ ,  $\beta = 3$ ,  $\omega = 0.3$ . The average population strategy for every task present in the environment is represented in red.

## References

1. Robson, S.K.; Traniello, J.F. Division of labor in complex societies: A new age of conceptual expansion and integrative analysis. *Behav. Ecol. Sociobiol.* **2016**, *70*, 995–998. [[CrossRef](#)]
2. Beshers, S.N.; Fewell, J.H. Models of division of labor in social insects. *Annu. Rev. Entomol.* **2001**, *46*, 413–440. [[CrossRef](#)] [[PubMed](#)]
3. Jeanson, R.; Weidenmüller, A. Interindividual variability in social insects—proximate causes and ultimate consequences. *Biol. Rev.* **2014**, *89*, 671–687. [[CrossRef](#)] [[PubMed](#)]
4. Wilson, R. *Game-Theoretic Analysis of Trading Processes*; Stanford Univ Ca Inst for Mathematical Studies in the Social Sciences: Stanford, CA, USA, 1985.
5. Traniello, J.F.; Rosengaus, R.B. Ecology, evolution and division of labour in social insects. *Anim. Behav.* **1997**, *53*, 209–213. [[CrossRef](#)]
6. Hölldobler, B.; Wilson, E.O. *The Ants*; Harvard University Press: Cambridge, MA, USA, 1990.

7. Fewell, J.H.; Bertram, S.M. Division of labor in a dynamic environment: Response by honeybees (*Apis mellifera*) to graded changes in colony pollen stores. *Behav. Ecol. Sociobiol.* **1999**, *46*, 171–179. [[CrossRef](#)]
8. Gazda, S.K.; Connor, R.C.; Edgar, R.K. Cox F. A division of labour with role specialization in group-hunting bottlenose dolphins (*Tursiops truncatus*) off Cedar Key, Florida. *Proc. R. Soc. Biol. Sci.* **2005**, *272*, 135–140. [[CrossRef](#)]
9. Rieger, N.S.; Stanton, E.H.; Marler, C.A. Division of labour in territorial defence and pup retrieval by pair-bonded California mice, *Peromyscus californicus*. *Anim. Behav.* **2019**, *156*, 67–78. [[CrossRef](#)]
10. Siefert, P.; Buling, N.; Grünwald, B. Honey bee behaviours within the hive: Insights from long-term video analysis. *PLoS ONE* **2021**, *16*, e0247323. [[CrossRef](#)]
11. Charbonneau, D.; Blonder, B.; Dornhaus, A. Social insects: A model system for network dynamics. *Temporal Netw.* **2013**, *2*, 217–244.
12. Grimaldi, D.; Engel, M.S.; Engel, M.S. *Evolution of the Insects*; Cambridge University Press: Cambridge, MA, USA, 2005.
13. Charbonneau, D.; Dornhaus, A. When doing nothing is something. How task allocation strategies compromise between flexibility, efficiency, and inactive agents. *J. Bioecon.* **2015**, *17*, 217–242. [[CrossRef](#)]
14. Gordon, D.M. From division of labor to the collective behavior of social insects. *Behav. Ecol. Sociobiol.* **2016**, *70*, 1101–1108. [[CrossRef](#)]
15. Hunt, G.J.; Amdam, G.V.; Schlipalius, D.; Emore, C.; Sardesai, N.; Williams, C.E.; Rueppell, O.; Guzmán-Novoa, E.; Arechavaleta-Velasco, M.; Chandra, S.; et al. Behavioral genomics of honeybee foraging and nest defense. *Naturwissenschaften* **2007**, *94*, 247–267. [[CrossRef](#)]
16. Scheiner, R.; Page, R.E.; Erber, J. Sucrose responsiveness and behavioral plasticity in honey bees (*Apis mellifera*). *Apidologie* **2004**, *35*, 133–42. [[CrossRef](#)]
17. Gordon, D.M. The organization of work in social insect colonies. *Complexity* **2002**, *8*, 43–46. [[CrossRef](#)]
18. Fewell, J.H.; Harrison, J.F. Scaling of work and energy use in social insect colonies. *Behav. Ecol. Sociobiol.* **2016**, *70*, 1047–1061. [[CrossRef](#)]
19. Mersch, D.P. The social mirror for division of labor: What network topology and dynamics can teach us about organization of work in insect societies. *Behav. Ecol. Sociobiol.* **2016**, *70*, 1087–1099. [[CrossRef](#)]
20. Bonabeau, E.; Theraulaz, G.; Deneubourg, J.L. Quantitative study of the fixed threshold model for the regulation of division of labour in insect societies. *Proc. R. Soc. Lond. Ser. Biol. Sci.* **1996**, *263*, 1565–1569.
21. Kang, Y.; Theraulaz, G. Dynamical models of task organization in social insect colonies. *Bull. Math. Biol.* **2016**, *78*, 879–915. [[CrossRef](#)] [[PubMed](#)]
22. Cook, C.N.; Breed, M.D. Social context influences the initiation and threshold of thermoregulatory behaviour in honeybees. *Anim. Behav.* **2013**, *86*, 323–329. [[CrossRef](#)]
23. Greene, M.J.; Gordon, D.M. Interaction rate informs harvester ant task decisions. *Behav. Ecol.* **2007**, *18*, 451–455. [[CrossRef](#)]
24. Duarte, A.; Weissling, F.J.; Pen, I.; Keller, L. An evolutionary perspective on self-organized division of labor in social insects. *Annu. Rev. Ecol. Evol. Syst.* **2011**, *42*, 91–110. [[CrossRef](#)]
25. Couzin, I.D. Collective cognition in animal groups. *Trends Cogn. Sci.* **2009**, *13*, 36–43. [[CrossRef](#)]
26. Leadbeater, E.; Chittka, L. Social learning in insects—From miniature brains to consensus building. *Curr. Biol.* **2007**, *17*, R703–R713. [[CrossRef](#)]
27. Camazine, S.; Crailsheim, K.; Hrassnigg, N.; Robinson, G.E.; Leonhard, B.; Kropounigg, H. Protein trophallaxis and the regulation of pollen foraging by honey bees (*Apis mellifera* L.). *Apidologie* **1998**, *29*, 113–126. [[CrossRef](#)]
28. Worden, B.D.; Papaj, D.R. Flower choice copying in bumblebees. *Biol. Lett.* **2005**, *22*, 504–507. [[CrossRef](#)]
29. Leadbeater, E.; Chittka, L. Social transmission of nectar-robbing behaviour in bumble-bees. *Proc. R. Soc. B Biol. Sci.* **2008**, *275*, 1669–1674. [[CrossRef](#)]
30. Grueter, C.; Leadbeater, E. Insights from insects about adaptive social information use. *Trends Ecol. Evol.* **2014**, *29*, 177–184. [[CrossRef](#)]
31. Jones, P.L.; Ryan, M.J.; Chittka, L. The influence of past experience with flower reward quality on social learning in bumblebees. *Anim. Behav.* **2015**, *101*, 11–18. [[CrossRef](#)]
32. Czaczkes, T.J.; Grüter, C.; Ratnieks, F.L. Trail pheromones: An integrative view of their role in social insect colony organization. *Annu. Rev. Entomol.* **2015**, *60*, 581–599. [[CrossRef](#)]
33. Grüter, C.; Farina, W.M. The honeybee waggle dance: Can we follow the steps? *Trends Ecol. Evol.* **2009**, *24*, 242–247. [[CrossRef](#)]
34. Riley, J.R.; Greggers, U.; Smith, A.D.; Reynolds, D.R.; Menzel, R. The flight paths of honeybees recruited by the waggle dance. *Nature* **2005**, *435*, 205–207. [[CrossRef](#)]
35. Khajehnejad, M.; García, J.; Meyer, B. Explaining workers' inactivity in social colonies from first principles. *J. R. Soc. Interface* **2023**, *20*, 20220808. [[CrossRef](#)]
36. Smolla, M.; Alem, S.; Chittka, L.; Shultz, S. Copy-when-uncertain: Bumblebees rely on social information when rewards are highly variable. *Biol. Lett.* **2016**, *12*, 20160188. [[CrossRef](#)]
37. Geritz, S.A.; Metz, J.A.; Kisdi, É.; Meszéna, G. Dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett.* **1997**, *78*, 2024. [[CrossRef](#)]
38. Geritz, S.A.; Kisdi, E.; Meszéna, G.; Metz, J.A. Evolutionary singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* **1998**, *12*, 35–57. [[CrossRef](#)]

39. Doebeli, M.; Hauert, C.; Killingback, T. The evolutionary origin of cooperators and defectors. *Science* **2004**, *306*, 859–862. [CrossRef]
40. Smith, J.M. *Evolution and the Theory of Games*; Cambridge University Press: Cambridge, MA, USA, 1982.
41. Weibull, J.W. *Evolutionary Game Theory*; MIT Press: Cambridge, MA, USA, 1997.
42. Izquierdo, L.R.; Izquierdo, S.S.; Vega-Redondo, F. Learning and evolutionary game theory. *Encycl. Sci. Learn.* **2012**, *36*, 1782–1788.
43. McNamara, J.M. Towards a richer evolutionary game theory. *J. R. Soc. Interface* **2013**, *10*, 20130544. [CrossRef] [PubMed]
44. Ravary, F.; Lecoutey, E.; Kaminski, G.; Châline, N.; Jaisson, P. Individual experience alone can generate lasting division of labor in ants. *Curr. Biol.* **2007**, *17*, 1308–1312. [CrossRef]
45. Chittka, L.; Müller, H. Learning, specialization, efficiency and task allocation in social insects. *Commun. Integr. Biol.* **2009**, *2*, 151–154. [CrossRef]
46. Rendell, L.; Boyd, R.; Cownden, D.; Enquist, M.; Eriksson, K.; Feldman, M.W.; Fogarty, L.; Ghirl, A.S.; Lillicrap, T.; Lal, K.N. Why copy others? Insights from the social learning strategies tournament. *Science* **2010**, *328*, 208–213. [CrossRef]
47. Diwold, K.; Merkle, D.; Middendorf, M. Adapting to dynamic environments: Polyethism in response threshold models for social insects. *Adv. Complex Syst.* **2009**, *12*, 327–346. [CrossRef]
48. Bonabeau, E.; Sobkowski, A.; Theraulaz, G.; Deneubourg, J.L. Adaptive Task Allocation Inspired by a Model of Division of Labor in Social Insects. *InBCEC* **1997**, *36*–45.
49. Duarte, A.; Pen, I.; Keller, L.; Weissing, F.J. Evolution of self-organized division of labor in a response threshold model. *Behav. Ecol. Sociobiol.* **2012**, *66*, 947–957. [CrossRef] [PubMed]
50. Theraulaz, G.; Bonabeau, E.; Deneubourg, J.N. Response threshold reinforcements and division of labour in insect societies. *Proceedings of the Royal Society of London. Ser. B Biol. Sci.* **1998**, *265*, 327–332. [CrossRef]
51. Seel, N.M. (Ed.) *Encyclopedia of the Sciences of Learning*; Springer Science & Business Media: Berlin, Germany, 2011.
52. Sandholm, W.H. *Population Games and Evolutionary Dynamics*; MIT Press: Cambridge, MA, USA, 2010.
53. Cross, J.G. A stochastic learning model of economic behavior. *Q. J. Econ.* **1973**, *87*, 239–266. [CrossRef]
54. Charbonneau, D.; Poff, C.; Nguyen, H.; Shin, M.C.; Kierstead, K.; Dornhaus, A. Who are the “lazy” ants? The function of inactivity in social insects and a possible role of constraint: Inactive ants are corpulent and may be young and/or selfish. *Integr. Comp. Biol.* **2017**, *57*, 649–667. [CrossRef]
55. Herbers, J.M. Time resources and laziness in animals. *Oecologia* **1981**, *49*, 252–262. [CrossRef]
56. Piezon, S.L. *Social Loafing and Free Riding in Online Learning Groups*; The Florida State University: Tallahassee, FL, USA, 2011.
57. Grossman, S.J.; Hart, O.D. Takeover bids, the free-rider problem, and the theory of the corporation. *Bell J. Econ.* **1980**, *11*, 42–64. [CrossRef]
58. Heinsohn, R.; Packer, C. Complex cooperative strategies in group-territorial African lions. *Science* **1995**, *269*, 1260–1262. [CrossRef]
59. Erev, I.; Roth, A.E. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* **1998**, *88*, 848–881.
60. Oldroyd, B.P.; Fewell, J.H. Genetic diversity promotes homeostasis in insect colonies. *Trends Ecol. Evol.* **2007**, *22*, 408–413. [CrossRef] [PubMed]
61. Oster, G.F.; Wilson, E.O. *Caste and Ecology in the Social Insects*; Princeton University Press: Princeton, NJ, USA, 1978.
62. Robinson, G.E. Regulation of honey bee age polyethism by juvenile hormone. *Behav. Ecol. Sociobiol.* **1987**, *20*, 329–338. [CrossRef]
63. Börgers, T.; Sarin, R. Learning through reinforcement and replicator dynamics. *J. Econ. Theory* **1997**, *77*, 1–4. [CrossRef]
64. Lahkar, R.; Seymour, R.M. Reinforcement learning in population games. *Games Econ. Behav.* **2013**, *80*, 10–38. [CrossRef]
65. Li, L.; McCann, J.; Faloutsos, C.; Pollard, N.S. Laziness Is a Virtue: Motion Stitching Using Effort Minimization. In Proceedings of the Eurographics ‘08 (Short Papers), Crete, Greece, 14–18 April 2008; pp. 87–89.
66. Charbonneau, D.; Sasaki, T.; Dornhaus, A. Who needs ‘lazy’ workers? Inactive workers act as a ‘reserve’ labor force replacing active workers, but inactive workers are not replaced when they are removed. *PLoS ONE* **2017**, *12*, e0184074. [CrossRef]
67. Lindauer, M. A contribution to the question of the division of labor in the bee colony. *J. Comp. Physiol.* **1952**, *34*, 299–345.
68. Michener, C.D. Reproductive efficiency in relation to colony size in hymenopterous societies. *Insectes Sociaux* **1964**, *11*, 317–341. [CrossRef]
69. Feng, T.; Charbonneau, D.; Qiu, Z.; Kang, Y. Dynamics of task allocation in social insect colonies: Scaling effects of colony size versus work activities. *J. Math. Biol.* **2021**, *82*, 1–53. [CrossRef]
70. Cirelli, C.; Tononi, G. Is sleep essential? *PLoS Biol.* **2008**, *6*, e216. [CrossRef]
71. Siegel, J.M. Do all animals sleep? *Trends Neurosci.* **2008**, *31*, 208–213. [CrossRef] [PubMed]
72. Jeanne, R.L. The organization of work in *Polybia occidentalis*: Costs and benefits of specialization in a social wasp. *Behav. Ecol. Sociobiol.* **1986**, *19*, 333–341. [CrossRef]
73. Johnson, B.R. Global information sampling in the honey bee. *Naturwissenschaften* **2008**, *95*, 523–530. [CrossRef] [PubMed]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.