



# Grounded predictions of teamwork as a one-shot game: A multiagent multi-armed bandits approach

Alejandra López de Aberasturi Gómez <sup>ID</sup>\*, Carles Sierra <sup>ID</sup>, Jordi Sabater-Mir <sup>ID</sup>

Artificial Intelligence Research Institute, Barcelona, IIIA-CSIC, Campus UAB Carrer de Can Planas, Zona 2, Barcelona, 08193, Spain

## ARTICLE INFO

### Keywords:

Cooperative AI  
Multiagent multi-armed bandits  
Group productivity theory  
Aggregative games

## ABSTRACT

Humans possess innate collaborative capacities. However, effective teamwork often remains challenging. This study delves into the feasibility of collaboration within teams of rational, self-interested agents who engage in teamwork without the obligation to contribute. Drawing from psychological and game theoretical frameworks, we formalise teamwork as a one-shot aggregative game, integrating insights from Steiner's theory of group productivity. We characterise this novel game's Nash equilibria and propose a multiagent multi-armed bandit system that learns to converge to approximations of such equilibria. Our research contributes value to the areas of game theory and multiagent systems, paving the way for a better understanding of voluntary collaborative dynamics. We examine how team heterogeneity, task typology, and assessment difficulty influence agents' strategies and resulting teamwork outcomes. Finally, we empirically study the behaviour of work teams under incentive systems that defy analytical treatment. Our agents demonstrate human-like behaviour patterns, corroborating findings from social psychology research.

## 1. Introduction

Cooperation and alliances have long been integral to human prosperity, and remain vital in today's world [1]. While cooperation is not unique to humans, our species stands out for its ability to apply teamwork skills across diverse domains and to conceptualise teamwork itself [2]. Here, we understand teamwork as defined by The American Heritage Dictionary [3]: *Cooperative effort by the members of a group or team to achieve a common goal*. Human cooperation extends beyond small groups to encompass large numbers of unrelated individuals engaging in various forms of collective action, including activities that involve personal costs for the benefit of others, such as contributing to public goods [4]. This kind of collective action is essential for addressing modern challenges outlined in sustainable development objectives, as stated by The United Nations General Assembly [5]. However, while collaboration comes naturally to us, achieving effective teamwork is still challenging.

From a mathematical perspective, teamwork has been extensively modelled within cooperative game theory, which assumes that all team members share a common reward and there is no conflict among them [6,7]. Solution concepts like the Shapley Value aim to fairly divide the total utility created by the team [8–11]. This approach works well in closely-knit teams where cooperation is clearly beneficial. However, in typical human teamwork scenarios, securing binding collaboration agreements is often unattainable, leaving room for defection and calling for the application of non-cooperative game theory.

\* Corresponding author.

E-mail addresses: [alejandra@iia.csic.es](mailto:alejandra@iia.csic.es) (A. López de Aberasturi Gómez), [sierra@iia.csic.es](mailto:sierra@iia.csic.es) (C. Sierra), [jsabater@iia.csic.es](mailto:jsabater@iia.csic.es) (J. Sabater-Mir).

On the other hand, computer science, particularly in the field of Multiagent Systems (MAS), has extensively studied teamwork and how to design multiagent systems that cooperate effectively [2,12,13]. Specifically, considerable attention has been directed towards team composition<sup>1</sup> and team formation<sup>2</sup> within groups of cooperative, self-interested agents [14]. Although the objectives necessitating cooperation vary among algorithms (e.g., minimising overall costs, maximising social utility, or enhancing outcome quality), the prevalent approach assumes that agents need cooperation to succeed individually, making intentionally defective strategies unviable. This approach primarily seeks to determine how to achieve maximum performance from the agent team, whether in terms of maximising benefits or minimising costs. In this sense, much of the focus has been on prescriptive approaches: *How can we design each agent to ensure effective cooperation? How should these agents work together to solve problems?* [13]. In short, the focus of most work on multiagent systems has been addressing issues associated with *self-interested* agents who require cooperation to achieve individual success, and resolving the resulting conflicts of interest to enable this cooperation. Surprisingly, the problem of predicting the performance of human teams, or more broadly, performance in teams without enforced cooperation, has been largely overlooked. When addressing this challenge, a descriptive approach, rather than a prescriptive one, would be more suitable: *What social effects arise when each (self-interested) agent in a society adopts a particular, not necessarily cooperative, policy?* The scarce efforts in computer science to tackle this issue have often lacked an integrative, cross-disciplinary understanding of progress made in fields like economics, game theory, and social psychology. For instance, Andrejczuk et al. [15] highlighted that in MAS models of human team formation and performance, there's often an unrealistic assumption that agents have complete knowledge of their environment and teammates. Plus, factors that are crucial according to psychology and economics, such as time constraints, task type and team heterogeneity, are often overlooked in MAS.

For instance, when collaborating in a team, individuals consider their teammates' expertise and their apparent commitment to the group's success to determine how much responsibility to delegate to them. Moreover, achieving a good outcome in a domain where one lacks proficiency often demands significantly more effort compared to areas where one has greater aptitude. The expected rewards or incentives for successful group performance, as well as potential punishments for unsatisfactory outcomes, also play a key role in shaping effort allocation. Time constraints are another critical factor; the amount of time one is willing to dedicate to a month-long project will naturally differ from the effort invested in a task that lasts only two days. Furthermore, the nature of the task itself can impose constraints on how much individual effort translates into improved group outcomes. For example, a cellist in a string quartet may spend countless hours practising their part, yet if the two violinists and the violist have not rehearsed adequately, the ensemble's performance will lack harmony. In such cases, the cellist may (rightly) conclude that their efforts were in vain, as the very structure of quartet performance limits the group's success to that of its weakest member. However, this limitation is not universal to all tasks. Certain teamwork settings allow a group to succeed as long as at least one member achieves the desired outcome. These tasks often involve intellectual challenges or decision-making scenarios where the group must collectively endorse a single contribution, such as finding the exit to a maze or solving a problem. Finally, there are tasks in which every unit of effort invested by team members is directly translatable into tangible benefits. For instance, in the context of a volunteer emergency response team addressing the aftermath of a natural disaster, every shovelful of mud cleared, every fallen tree removed, or every brush stroke applied to repaint a wall constitutes a valuable contribution that can be added to the public good and that benefits the affected community.

In this paper, we address this gap in the literature by modelling human teamwork through a descriptive approach. We move away from the assumptions of complete agent knowledge and enforced cooperation, instead integrating insights from psychology and game theory to answer the question of how to learn theoretically grounded predictions of team performance.

This research question has two main components: Firstly, our goal is to identify the Nash Equilibria (NE) in a non-cooperative game that models teamwork. Secondly, we aim to propose a learning framework where agents converge to the teamwork outcomes predicted by the NE of such a game.

As per the second aspect, the topic of learning in games studies how less than fully rational players grope for optimality over time, as well as understand possible barriers to reaching equilibrium.

Specifically, although reinforcement learning (RL) agents do not generally converge to a Nash equilibrium, they can use value-based learning techniques to approximate Nash distributions when employing smooth best responses—that is, joint mixed strategies where each player adopts a smooth best response to rewards shaped by their opponents' actions [16].

The architecture of RL agents naturally lends itself to representing beliefs, preferences, intentions, and actions [17]. When combined in a multiagent system, the interactive learning of RL agents facilitates the emergence of complex behaviours [18–20]. Thus, Multiagent Reinforcement Learning (MARL) provides a suitable framework to explore the emergence of social dynamics during teamwork through a descriptive approach. Traditionally, teams in MARL have been fully cooperative, with agents sharing a common reward [21]. However, the field is evolving towards more realistic multiagent teams, including mixed-motive scenarios and social dilemmas beyond zero-sum games [22], more detailed modelling of agents' motivations [19,23–27], and innovative multiagent negotiation techniques [28].

These efforts have contributed to the rise of a new research stream called “Cooperative AI” [29]. The goal of Cooperative AI is to promote interdisciplinary dialogue with natural, social, and behavioural sciences. Cooperative AI leverages recent advances in artificial intelligence, including reinforcement learning, to study various aspects of cooperation and solve problems arising in systems of (human and/or machine) agents. For example, Gemp et al. [30] framed learning incentives as an inefficiency problem within game outcomes—i.e., the extent to which game results are suboptimal both from individual agent perspectives and at the collective level.

<sup>1</sup> Andrejczuk et al. [14] define team composition as the process of deciding which agents should be part of a team.

<sup>2</sup> Conversely, Andrejczuk et al. [14] define team formation as the process undertaken by agents to learn to work together in a team.

Consequently, their deep-learning agents developed policies that enhanced outcomes for both individual agents and the group overall in environments characterised by social dilemmas. However, despite the development of more realistic multiagent teams in MARL, a Cooperative AI approach has not yet been used to model teamwork without presuming enforced or incentivised cooperation. We see an opportunity to improve the modelling of human teamwork by harnessing MARL frameworks to represent selfish, interactive learning agents who are not compelled to cooperate, drawing on insights from psychology.

### 1.1. Our contribution

Our research addresses a pivotal question: How can we effectively represent **(1) a teamwork setting without enforced cooperation** and **(2) a team of reward-maximising agents** to learn theoretically grounded predictions of team performance? To answer this, we model teamwork as a mixed-motive game, drawing inspiration from social psychology research on human behaviour in teams. We characterise the game's equilibria, which predict teamwork outcomes. Building on previous research in learning in games, we propose a Multiagent Multi-Armed Bandit (MA-MAB) system to learn approximations of the Nash Equilibria (NE) of the game. Our methodology is grounded in aggregative games as the foundational framework. As will be explained shortly, aggregative games are a subset of pure-strategy non-cooperative games that express payoffs for each player as a function of their strategy and an aggregate representing all interactions in the game. This condensed payoff representation overcomes the complexities associated with an increasing number of players [31]. Although previous research has used bandits to play games such as Cournot oligopoly one-shot games [32], zero-sum games [16], and repeated Stackelberg games [33], to the best of our knowledge, this is the first instance where aggregative games have been explicitly combined with Multiagent Multi-Armed Bandit Models. This work aims to contribute to Cooperative AI [29], both theoretically and practically:

- We introduce *teamwork games*, a novel aggregative game that explicitly models teamwork and considers the evaluation of teamwork outcomes as part of the players' utility function. Our work builds on the most popular research stream in interdependence theory [34] on teamwork: Steiner's [35] *theory of group productivity*. We identify the branch of game theory that allows us to mathematically formulate the main ingredients of Steiner's theory: *general public good games* [36,37]. Extending this framework, we incorporate variables that have been previously overlooked, including the weights of team members' opinions within the team, task complexity, individual expertise, preferences, and time constraints.
- We characterise the NE of this novel game, and find that they exhibit variations of the equilibria observed in general public good games [36,37]. These variations are influenced by elements unique to our model, such as the teamwork evaluation function and player expertise.
- We propose a multiagent framework in which agents learn strategic behaviour that approximates the NE of the game. Building on the results in learning in games by Leslie and Collins [16], we conceptualise the allocation of individual work contributions to the team task as a multi-armed bandit problem with Boltzmann action selection.
- We validate the proposed MA-MAB system by empirically proving the convergence of the learned strategies towards the approximated NE of the game. In over 140 experiments, the model achieved a near-perfect fit ( $\xi^2 = 0.992$ ) between the equilibrium aggregate of team productivity and the actual productivity to which the MA-MAB model converged. Additionally, we analyse the impact of isolated variables on team productivity when the bandits face teamwork scenarios:
  - When examining how agent productivity varies with different types of tasks, our findings align with Steiner's hypotheses on team composition and productivity. Specifically, tasks where the weakest team members significantly impact the outcome are better performed by homogeneous teams with similar expertise levels. In additive tasks, where each member's contribution adds up to the outcome, team heterogeneity does not affect productivity. Conversely, tasks where the strongest members have the most influence are better handled by teams with diverse expertise levels.
  - When we analyse the effects of increasing the difficulty of the teamwork evaluation function, we find that productivity rises with task difficulty, especially when passing to the next evaluation level is achievable. However, this trend halts once the task requirements exceed the team's capacity. At this point, agents either refrain from increasing their contribution or reduce it.
  - We find that agents with lower expertise are more influenced by changes in task type, while highly skilled agents are more responsive to changes in the teamwork evaluation function than to the task's nature.
- After validating the MA-MAB framework, we expand our analysis beyond analytically solvable games and empirically study the policies of our agents after learning with a discontinuous evaluation function, resembling a binary pass/fail assessment.

The remainder of the article is structured as follows. Section 2 explores the theoretical frameworks related to teamwork and decision-making from the perspectives of game theory, social psychology, and multiagent systems. Section 3 introduces key concepts from aggregative games, forming the foundation of our mathematical framework for investigating teamwork. Section 4 builds on this framework to establish a model of human teamwork. Section 5 examines the equilibria in teamwork games, highlighting the effects of different parameter regimes on the team outcome. Section 6 conceptualises the decision-making process of team members as a multiagent multi-armed bandit problem, detailing the learning dynamics involved. Section 7 presents our experimental setup and discusses the results, focusing on team productivity, individual strategies, and simulations beyond traditional aggregative game theory. Finally, Section 8 draws conclusions, discusses limitations, and outlines potential future work.

## 2. Preliminaries

In this section, we explore the theoretical frameworks related to teamwork or decision-making from the perspectives of the three main areas integrated in this paper: game theory, social psychology, and multiagent systems.

### 2.1. Aggregative games and the public goods dilemma

From a mathematical perspective, numerous efforts have been made to model teamwork, particularly within cooperative game theory. Cooperative game theory approaches assume that all team members share a common reward, and thus there is no potential for conflict between them. Solution concepts in cooperative game theory, such as the Shapley Value, have been developed to be “fair” divisions of the total utility created by the team or coalition [21,38–40].

While this approach may be effective for tightly-knit teams where the benefits of cooperation (or the drawbacks of non-cooperation) are evident, it often falls short in typical human teamwork settings. In these scenarios, it’s not feasible to establish binding agreements that ensure the cooperation of all team members. Consequently, it becomes necessary to acknowledge the possibility that team members may shirk their collective responsibilities and seek to maximise their own utility by free-riding on the efforts of others. This perspective on teamwork is further explored through mixed games, notably in the context of the voluntary provision of public goods game<sup>3</sup> [41–43]. Indeed, teamwork can be viewed as a public good, presenting members with a dilemma between contributing to the collective task or exploiting others’ efforts (social loafing). Despite the potential benefits of cooperation, individuals may have varying preferences regarding the provision of the common good, and practical constraints often fail to prevent free-rider behaviours [44]. In this paper, we adopt this perspective of teamwork as a mixed-motive game.

Aggregative games, a subset of pure-strategy non-cooperative games, express payoffs for each player as a function of their strategy and an aggregate representing all interactions in the game. This approach simplifies analysis as it avoids the complexities associated with an increasing number of players [31]. By aggregating team interactions, analysts can simplify calculations, eliminating the need to consider individual strategies to compute the best response. Notably, some longstanding economic models exhibit this aggregative structure, including the voluntary provision of public goods game [45], which is the focus of this paper.

In the literature on aggregative games, the emphasis lies in computing Nash Equilibria (NE) and understanding how NE are influenced by changes in exogenous parameters (an analysis called comparative statics), along with addressing their existence [31,46]. The antecedents of using aggregative games as an analytical tool for the voluntary contribution to public goods are presented by Cornes and Hartley [37]. Their work adapted the canonical model of public goods proposed by Bergstrom et al. [47]—which considers public goods as the unweighted sum of all individual contributions—to the aggregative games formalism.

However, as early as 1983, Hirschleifer [48] had already suggested two types of public goods—best-shot (where the best individual performance determines team productivity) and weakest-link (where the poorest individual performance determines team productivity)—that are not captured by summation technology but hold empirical significance. Cornes [49] and Cornes and Hartley [36] address these observations by exploring games where the total public good level is generated by individual contributions based on a more general, not necessarily summative, production technology. For further insights, interested readers may consult works such as [50] and [51].

Using aggregative games, Jensen [52] examines the effect of diverse incentive systems on team performance within heterogeneous teams characterised by varying skill sets and expertise levels. Although Jensen’s approach shares similarities with ours, it differs in that his analysis posits that contributions from *cognitively* similar members in a team are additive. The non-linearities in their model arise from the complexity of coordinating efforts among members of different cognitive categories<sup>4</sup> within diverse teams, rather than variations in task type.

### 2.2. Social psychological theories of group productivity

In the field of psychology, social dilemmas, especially the public goods dilemma, have been thoroughly examined through interdependence theory [53,54]. The theory of interdependence, developed by Lewin, views groups as dynamic entities shaped by the interdependence among their members [34]. Understanding teamwork as an instance of the public goods dilemma, we elaborate on a renowned theory of team effectiveness aligned with interdependence theory: Steiner’s model [35]. While we acknowledge that team effectiveness encompasses more than mere productivity [55], in this paper, we focus on productivity as the metric to evaluate teamwork [56,57]. Plus, we focus our attention on the internal variables within a group that affect and are best predictors of group productivity (*i.e.* closed group paradigm [58]). Steiner’s model identifies four key elements within a group that influence productivity:

1. **Individuals’ profiles:** Individuals contribute various resources to the task, ranging from skills and expertise levels to effort and personality traits.
2. **Task Type:** Steiner categorised tasks based on how individual contributions combine to shape the group’s performance (see Table 1); which parallels Cornes and Hartley’s observation that the nature of the public good dictates how the team resources

<sup>3</sup> In this basic version of the game, participants privately decide how many tokens to contribute to a communal pool. These tokens are aggregated, resulting in a collective outcome known as the “public good”. The benefits from this public good are evenly distributed among all players, while each participant retains the tokens they choose not to contribute.

<sup>4</sup> Jensen [52] defines a cognitive category as a specific combination of approaches, perspectives, and abilities.

**Table 1**  
Steiner's task taxonomy based on individual contributions and group outcome [35].

Task Type	Characteristics	Example
Additive	Group success results from the sum of individual efforts, with each member's contribution adding to the collective outcome.	A group project where each member researches and writes a specific section, contributing to a comprehensive report.
Conjunctive	Group performance is determined by the contribution of the group member who performs most poorly.	In times of disaster, community survival depends on each person fulfilling their duty. Heroism emerges when any individual can become the weak link in dire circumstances [48,59].
Disjunctive	Group performance is determined by the contribution of the member who performs the most.	Imagine multiple anti-missile batteries defending a city against a single incoming nuclear-armed ballistic missile. Success hinges on whether the best shot can destroy the incoming threat [48].
Discretionary	Members' contributions are combined through weighted averaging, often in tasks involving judgment.	When tasked with estimating the temperature of a room, group members engage in discussion, and their final consensus is usually an implicit weighted average of their opinions, where the weights signify the importance of each member's opinion within the team.

should be combined to achieve a favourable collective outcome. For instance, in activities like tug-of-war, each participant's physical effort directly contributes to the total force exerted. However, dynamics are different in scenarios like marching soldiers or an orchestra, where the slowest or least proficient member can determine the overall performance. Steiner's task taxonomy is still relevant, explaining how the nature of a collective task determines the integration of individual contributions into the final outcome. Additionally, he distinguished between **unitary** tasks (tasks that cannot be subdivided into smaller subtasks and require a single skill for completion) and **divisible** tasks (tasks that are readily divided into subtasks, each performable by a different individual, usually involving multiple skills). In this paper, we focus solely on unitary tasks.

3. **Group Interaction Processes:** Team members' preferences and intentions drive their action choices. Introducing an evaluation system that assesses group production will most possibly impact those preferences, reshaping the interdependencies among members and thereby affecting productivity. Steiner [35] identified three main types of payoff systems: *promotive* or cooperative (each team member receives a high payoff when their behaviour is highly beneficial to their partners and to themselves), *construent* or competitive (team members exert contrary effects upon one another's payoff), and *independent* (where each team member's payoff is unaffected by the rest of the team's actions). In this paper, we focus on the first paradigm, studying the impact of introducing incentives for cooperation during group evaluation.
4. **Group Outcome:** Individual decisions are influenced by the foreseen group's outcome, particularly whether the group's outcome will satisfy the individuals' preferences or not.

Steiner clearly distinguished between potential performance, representing a group's capabilities, and actual performance, reflecting its accomplishments. If somehow one knew a team's potential performance, its actual performance would be obtained by subtracting some "lost" processes from its potential performance. These processes refer to any factor hindering the group from reaching its full potential. For instance, Steiner labelled as *motivation loss* the situation when one or several members exert less effort than their potential in a collective task. One factor contributing to motivation loss is *social loafing*, a free-rider behaviour characterised by decreased individual contributions when working in a group compared to working alone, assuming others are also working [60–63]. Conversely, *social compensation* entails individuals putting in extra effort when anticipating lower performance from group members. This phenomenon arises from the desire to compensate for perceived skill or motivation gaps among colleagues to complete a task successfully. The literature on social compensation suggests that individuals are more likely to work harder in a group setting if 1) at least one group member anticipates insufficient effort from others for success, and 2) the task holds significant importance to those individuals [64]. Similar to the connection between social loafing in social psychology and free-riding in public goods dilemmas, many concepts emerging from the study of teamwork through a psychosocial lens closely align with established ideas in game theory. For example, the distinction made by Steiner between potential and actual performance has been described in algorithmic game theory as the "price of anarchy" [65]. This term is a technical measure that compares the welfare achievable through optimal cooperation with the welfare attained by self-interested agents in a particular equilibrium scenario.

### 2.3. Multiagent multi-armed bandits (MA-MAB) and decision-making

Human behaviour, particularly decision-making, can be likened to a search problem of actions that lead to favourable outcomes. Decision-makers must navigate the balance between exploring new options and exploiting known ones. Researchers often employ multi-armed bandit (MAB) models to depict human decision-making within exploration-exploitation tasks. These models represent decision-makers faced with multiple options, each associated with uncertain rewards [66,67]. The classical bandit problem simplifies decision-making into a one-move one-person game with random payoffs for each action. In a  $k$ -armed bandit problem some action payoffs remain unknown, and observing the payoff from one action provides no information about the distributions governing the



payoffs of other choices [68]. Specifically, each action (arm)  $a$  is associated with an expected reward if selected, denoted by  $q(a)$ , which represents the action's value.

Each agent maintains an estimate of the value of action  $a$  at time  $t$ , denoted as  $Q_t(a)$ . The value estimates are computed incrementally. Let  $Q_t$  denote the estimate after  $t - 1$  previous selections for a particular action. Given this estimate and the  $t^{\text{th}}$  reward  $R_t$ , the value estimate is updated using the following rule:

$$Q_{t+1} = Q_t + l_r \cdot (R_t - Q_t) \quad (1)$$

In this equation,  $l_r \in (0, 1]$  is the learning rate parameter and modulates the influence of new rewards on the estimated value. Higher learning rates imply greater adaptability to new information, while lower rates prioritise stability in prior estimates. We will pick a learning rate with a decay schedule that ensures asymptotic stochastic convergence, meeting the conditions:

$$\begin{cases} \sum_{t=0}^{\infty} l_r = \infty \\ \sum_{t=0}^{\infty} (l_r)^2 < \infty \end{cases}$$

Please note that while our work aims to model teamwork dynamics, it does not attempt to capture the specific process by which team members may learn strategic behaviour. Rather, we focus on comparing the multi-agent strategies to human behaviours, as discussed in the social psychology literature, only *post-training*. Consequently, the role of the Q-learners' learning rate was to ensure convergence rather than to emulate human learning processes. For readers interested in reinforcement models of human learning in strategic contexts, seminal works such as [69,70] offer further insight.

Those interested in alternative forms of the learning rate that more closely model human behaviour may refer to Child et al. [71], who propose an exponential decay learning rate to simulate decision-making grounded in heuristic reasoning [72].

Leslie and Collins [16] discuss how employing action selection methods based on smoothed best responses can guide a multiagent multi-armed bandit system toward a set of strategies approximating a Nash equilibrium. This approach aligns closely with the concept of quantal equilibria, first introduced by McKelvey and Palfrey [73]. Quantal equilibria extend the Nash equilibrium framework by incorporating smoothed best responses, where players' actions are probabilistic rather than deterministic. This modelling leverages standard statistical models for quantal choice [74], replacing the perfectly rational expectations equilibrium of Nash with a "noisy" rational expectations equilibrium that accounts for bounded rationality.

The quantal equilibrium framework has proven to be a powerful tool for capturing human decision-making processes, as demonstrated by studies such as Tumennasan [75], Kozitsina et al. [76], Goeree et al. [77], Ge and Godager [78]. In reinforcement learning, this approach is operationalised through the soft-max action selection method [68], which implements a specific form of quantal equilibrium known as the logit equilibrium [79].

Soft-max action selection calculates the probability of choosing an action based on its estimated value and a temperature parameter,  $\tau > 0$ , which controls the degree of exploration versus exploitation:

$$\mathbb{P}(a) = \frac{e^{Q_t(a)/(Q_t^{\max, \tau})}}{\sum_{a'} e^{Q_t(a')/(Q_t^{\max, \tau})}}$$

Here,  $Q_t^{\max}$  represents the value of the best-valued action(s) at time  $t$ . By smoothing the decision-making process, soft-max enables agents to explore suboptimal actions occasionally, ensuring that the system balances exploration and exploitation as it converges toward an equilibrium.

This action selection method implements smooth best responses. The use of smooth best responses implies that Nash Equilibria (NE) are no longer fixed points in the strategy space. Instead, the algorithm will converge to Nash distributions. Nash distributions are joint strategies where each agent plays smooth best responses to the rewards arising during interaction with others. According to Govindan et al. [80], NE of a game are approximated by Nash distributions for sufficiently low temperatures. Therefore, if  $\tau$  is sufficiently low and the agents interact for a sufficiently long period, the learning algorithm can be blindly applied in any game. If convergence occurs, a Nash distribution must have been reached [16].

### 3. Background: generalised aggregative games

In this section, we introduce key concepts from aggregative games that underpin one of our main contributions developing a theoretical framework to investigate teamwork formally. In this introductory section, we use intentionally broad and standard notation to ensure compatibility with most work in the field. However, as we delve into the formulation of our specific problem (Section 4), certain adaptations to the notation will be necessary to accommodate the nuances of our model. For a comprehensive review of aggregative games, the interested reader may refer to Jensen et al. [31] and Corchón [45].

Throughout, we will only consider games with a finite number of players, and a set of  $n$  players will be denoted by  $\mathcal{I} = \{1, \dots, n\}$ . Unless otherwise stated, we adhere to the following notation: each player  $i$  possesses an action or strategy  $d$ -dimensional set  $\mathcal{A}_i \subseteq \mathbb{R}^d$ , with  $a_i \in \mathcal{A}_i$  denoting a typical element. A joint action set is denoted  $\mathcal{A} \equiv \prod_{i=1}^n \mathcal{A}_i$ , with a joint action written as  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}$ . For any fixed player  $i$ , the vector of opponents' joint actions is  $\mathbf{a}_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n) \in \mathcal{A}_{-i}$ , with  $\mathcal{A}_{-i} \equiv \prod_{j \neq i} \mathcal{A}_j$ . We will consider that player  $i$ 's payoff function depends on their individual action and the vector of their opponent's joint actions:  $u_i : \mathcal{A}_i \times \mathcal{A}_{-i} \rightarrow \mathbb{R}$ . A (pure strategy Nash) equilibrium is a joint strategy  $\hat{\mathbf{a}} = (\hat{a}_i, \hat{\mathbf{a}}_{-i})$  such that:

$$u_i(\hat{a}_i, \hat{\mathbf{a}}_{-i}) \geq u_i(a_i, \hat{\mathbf{a}}_{-i}) \quad \forall a_i \in \mathcal{A}_i, i \in \mathcal{I} \quad (2)$$

Aggregative games are a subset of pure strategy non-cooperative games [31]. Their distinctive feature is that, given a player  $i$ , their utility function  $u_i$  can be simplified into a *reduced* utility function that depends on  $i$ 's action  $a_i$  and some finite aggregate of the set of joint actions of all players,  $G(\mathbf{a})$ :

$$\tilde{u}_i(a_i, G(\mathbf{a})) \equiv u_i(a_i, \mathbf{a}_{-i}) \quad (3)$$

The aggregation function  $G(\mathbf{a})$  determines how individual actions are combined to form a collective outcome. Depending on how the aggregate  $G(\mathbf{a})$  is computed, aggregative games can be classified into three categories: linearly aggregative, generalised aggregative, and quasi-aggregative games.

In linearly aggregative games, the aggregate  $G(\mathbf{a})$  is computed as a simple sum of the players' actions. Generalised aggregative games extend this framework by relaxing the strict requirement that the aggregation function be a sum. Instead, the aggregate is represented by an additively separable function, allowing for more complex relationships between individual strategies and their contribution to the group outcome. This flexibility is essential for capturing the nuanced interdependencies between players' actions in contexts like teamwork, where the impact of each player's effort may vary depending on the contributions of others.

Specifically, generalised aggregative games use transformations of individual actions, represented by functions  $h_i(g_i)$ , to compute the aggregate outcome. These transformations allow for scenarios where individual contributions amplify or diminish each other's effects depending on the nature of the task. For example, when tasks are *conjunctive* or *quasi-conjunctive*—where the group outcome is determined by the least effective member or a subset of the weakest contributors—the efforts of high-performing members may have limited influence beyond compensating for weaker contributions. Consider the case of a choir: even if most members sing in perfect harmony, a single member singing out of tune can detract significantly from the overall performance, potentially overshadowing the contributions of others. This highlights how, in such scenarios, the interdependencies between members become pronounced, and the performance of the “weakest link” carries disproportionate weight in determining the collective success. This illustrates the necessity of a more flexible framework, like that of generalised aggregative games, to adequately represent such task dynamics. In contrast, quasi-aggregative games further relax the structure by removing the strict positivity constraints on the aggregation functions  $H$  and  $h_i$  [31]. This allows for transformations of individual strategies that can include negative values or other complex forms, although such generality may not always be appropriate for contexts involving interdependent cooperation, such as teamwork scenarios.

As we will demonstrate in the following sections, generalised aggregative games align closely with the structure of our public goods model, where the group outcome is represented using a Constant Elasticity of Substitution (CES) function. This formulation captures the additive separability of players' contributions, expressed as:

$$G = H \left( \sum_{i=1}^n h_i(g_i) \right),$$

with  $H(z) = z^{1/\rho}$  and  $h_i(g_i) = \beta_i \cdot g_i^\rho$ . The CES function is particularly well-suited for modelling scenarios where the contributions of individual players interact in a non-linear way, capturing the interdependencies that arise in collective tasks.

**Definition 3.1** (*Generalised Aggregative Game* [31]). A non-cooperative game  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in \mathcal{I}}$  where  $\mathcal{A}_i \subseteq \mathbb{R}^d \forall i$  is termed generalised aggregative if:

1. There exists a social composition function  $G : \mathcal{A} \rightarrow \mathbb{R}$  (the *aggregator*) and a reduced utility function  $\tilde{u}_i : \mathcal{A}_i \times \mathcal{G} \rightarrow \mathbb{R}$ .
2. The aggregator  $G$  is an additively separable function:  $G(\mathbf{a}) = H(\sum_{i=1}^n h_i(a_i))$  where  $H : \mathbb{R} \rightarrow \mathbb{R}$  and  $h_i : \mathcal{A}_i \rightarrow \mathbb{R}$  are strictly increasing functions.
3. Each player's utility function  $u_i$  can be re-written using their reduced utility function  $\tilde{u}_i$ :

$$u_i(\mathbf{a}) = \tilde{u}_i(a_i, G(\mathbf{a}))$$

4. The values in the range of  $G$  are the set of possible aggregates:

$$\mathcal{G} \equiv \{G(\mathbf{a}) : \mathbf{a} \in \mathcal{A}\}$$

An aggregate is considered an *equilibrium aggregate* if  $\hat{\mathbf{a}}$  represents a pure strategy NE for the game.

During the computation of equilibria in generalised aggregative games, a crucial tool emerges: the *replacement correspondence*, closely linked to the concept of the *best-response correspondence*. Simply put, the best-response of a player  $i$  represents their optimal strategy in response to their opponents' strategies:

**Definition 3.2** (*Best-response correspondence* [31]). In a generalised aggregative game  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in \mathcal{I}}$ , player  $i$ 's best-response correspondence, denoted  $b_i : \mathcal{A}_{-i} \rightarrow 2^{\mathcal{A}_i}$ , is given by:

$$b_i(\mathbf{a}_{-i}) = \operatorname{argmax}_{a_i \in \mathcal{A}_i} \tilde{u}_i \left( a_i, H \left( \sum_{j=1}^n h_j(a_j) \right) \right) \quad (4)$$

By writing  $\sum_{j \neq i} h_j(a_j)$  as the sum of all players' strategies except for player  $i$ , we can define player  $i$ 's reduced best-response as a function of this sum:

$$\tilde{b}_i \left( \sum_{j \neq i} h_j(a_j) \right) = \operatorname{argmax}_{a_i \in \mathcal{A}_i} \tilde{u}_i \left( a_i, H \left( h_i(a_i) + \sum_{j \neq i} h_j(a_j) \right) \right) \quad (5)$$

By construction,  $b_i(\mathbf{a}_{-i}) = \tilde{b}_i(\sum_{j \neq i} h_j(a_j)) \forall \mathbf{a}_{-i} \in \mathcal{A}_{-i}$ . In other words, a player's best-reply is a function of the aggregate of the other players' actions.

The best-response function tells us the optimal action for a player considering their opponents' actions. Conversely, the replacement correspondence asks: given an outcome aggregate  $G$ , what actions could player  $i$  have taken that are consistent with  $G$  being the aggregate? In essence, it helps us identify the feasible actions for player  $i$  when we know the value of the outcome aggregate.

**Definition 3.3** (*Replacement Correspondence* [31,46]). Let  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in I}$  be a generalised aggregative game. The replacement correspondence for player  $i$ , denoted as  $r_i : \mathcal{G} \rightarrow 2^{\mathcal{A}_i}$ , is defined as:

$$r_i(G) \equiv \{a_i \in \mathcal{A}_i : a_i \in \tilde{b}_i(H^{-1}(G) - h_i(a_i))\} \quad (6)$$

Where  $G$  is some aggregate value and we have used that  $\sum_{j \neq i} h_j(a_j) = H^{-1}(G) - h_i(a_i)$ .

The aggregate replacement correspondence, denoted as  $R$ , expands on this concept. It computes the set of possible outcome aggregates resulting from each player's replacement correspondence over a given aggregate value. This concept proves instrumental in finding a Nash Equilibrium (NE) of an aggregative game, particularly as the number of players increases:

**Definition 3.4** (*Aggregate Replacement Correspondence* [46]). Let  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in I}$  be a generalised aggregative game. The aggregate replacement correspondence, denoted as  $R : \mathcal{G} \rightarrow 2^{\mathcal{G}}$ , is defined as:

$$R(G) \equiv \{G(\mathbf{a}) \in \mathcal{G} : a_i \in r_i(G) \forall i \in I\} \quad (7)$$

Jensen et al. [31] uses the aggregate replacement correspondence to define an equilibrium aggregate through the following proposition:

**Proposition 3.1.** *An aggregate  $\hat{G}$  is an equilibrium aggregate if and only if it is a fixed point of the aggregate replacement correspondence, expressed as:*

$$\hat{G} \in R(\hat{G}) \quad (8)$$

Thus, to investigate the equilibria of a generalised aggregative game, we can proceed as follows:

1. Compute  $R(G)$  as in Eq. (7).
2. Obtain the equilibrium aggregates  $\hat{G}$  as the fixed points of Eq. (8).
3. By Proposition 3.1, any fixed point  $\hat{G}$  determines a set  $\{\hat{\mathbf{a}}\}$  of (possibly unique) pure strategy Nash equilibria, which we compute by applying the set of replacement correspondences  $\{r_i\}_{i \in I}$  over  $\hat{G}$ :

$$\{\hat{\mathbf{a}}\} = \{(\hat{a}_1, \dots, \hat{a}_n) \in r_1(\hat{G}) \times r_2(\hat{G}) \times \dots \times r_n(\hat{G})\} \quad (9)$$

We will see that in some contexts, it is convenient to compute equilibria by means of the *share* correspondence:

**Definition 3.5** (*Share Correspondence*). Let  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in I}$  be a generalised aggregative game where player  $i$  has a replacement correspondence  $r_i(G)$ . Then, for any aggregate  $G > 0$ , the share correspondence  $s_i(G) : \mathcal{G} \rightarrow [0, 1]$  is given by:

$$s_i(G) = \left\{ \frac{h_i(a_i)}{H^{-1}(G)} : a_i \in r_i(G) \right\} \quad (10)$$

Rather than indicating the absolute contribution that player  $i$  makes to the teamwork outcome (as  $r_i$  does), this correspondence provides insights into the *relative* contribution by  $i$  concerning the total teamwork outcome. Whenever the replacement correspondence is a function, the share correspondence is also a function. According to Hartley and Cornes [81],  $i$ 's share function answers the following question:

*Given an aggregate value  $G$ , is there a proportion,  $b \in [0, 1]$ , such that, if  $b \cdot G$  were taken away from  $G$ , player  $i$ 's best response to the remaining quantity,  $(1 - b) \cdot G$ , would precisely compensate for the proportion  $b$ ?*

Suppose we can define a replacement and share function for every player in a game. Then, a strategy profile  $\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_n)$  is an equilibrium with equilibrium aggregate  $\hat{G}$  if and only if



$$S(\hat{G}) = \sum_{j=1}^N s_j(\hat{G}) = 1 \quad (11)$$

where  $S(\hat{G})$  is the aggregate share function [31].

Finally, it is pertinent to introduce the concept of normal good, given its relevance for this paper:

**Definition 3.6 (Normal Good).** A normal good is a type of good<sup>5</sup> whose demand increases as consumer income rises. Given a good  $x$  and a consumer's income  $y$ , the Income Elasticity of Demand  $\xi$  measures the sensitivity of demand for  $x$  to changes in  $y$ :

$$\xi = \frac{\Delta x / x}{\Delta y / y} \quad (12)$$

where  $\Delta x$  denotes the change in demand for  $x$ , and  $\Delta y$  denotes the change in income  $y$ .

We say that a good is normal if  $0 < \xi < 1$ .

In essence, normal goods, like food and clothing, demonstrate a positive correlation between demand and income.

#### 4. Teamwork games

Having introduced generalised aggregative games, we build on this framework to develop a model of human teamwork that captures the key elements of Steiner's theory of group tasks. Specifically, we model teamwork as a general public goods game, enabling us to represent the various task types outlined in the taxonomy presented in Section 1. For instance, we will demonstrate that certain parameter configurations in the model correspond to additive tasks. While pure additivity is rare outside laboratory settings, many collective physical activities (e.g., shovelling snow, pulling a rope, cleaning a room) can be broadly considered additive. Similarly, assembling a jigsaw puzzle as a group, or any task where individual contributions can be directly summed, also exemplifies additivity. Conjunctive tasks, where the group's success is determined by the performance of its least productive member, can also be captured within this model under a specific parameter regime. Steiner famously described the speed at which a team of mountain climbers can reach the top of a cliff as a quintessential example of such tasks. However, conjunctivity is not limited to extreme physical challenges; it is frequently observed in everyday scenarios. For instance, in some discussion groups, progress is constrained by the least capable participant, as group norms often require that every member fully understands a topic before moving on to the next. Similarly, elementary school teachers often lament that the pace of instruction is dictated by a few students who struggle to keep up with their peers. Indeed, the inherently conjunctive nature of classroom learning is one of the primary reasons behind the widespread practice of grouping students by ability [35]. Finally, disjunctive tasks are those where the group's success depends on selecting one member's performance to represent the collective effort. In such cases, the group product is essentially an individual product sanctioned by the group. These tasks are described as disjunctive because they involve an "either-or" decision process, where the group must accept one individual contribution as its own and reject all others. One member receives full "weight" in determining the group's outcome, while the rest are accorded no weight at all. Tasks are often made disjunctive when groups are required to choose between two or more distinctly different options as the final group product. For instance, this occurs when a group must answer "yes" or "no" to a question, select the correct solution to a problem, or choose between a few unspecified alternatives. In such cases, the only meaningful and permissible procedure is to adopt the preferred answer of one member as the group's outcome. An example of this, which we explore further in Subsection 4.2, is a team of students working together to solve an algebra problem.

In view of the above, building on the foundation of generalised aggregative games, we now narrow our focus to general public good games, as characterised by Cornes and Hartley [36]. This specialised framework allows us to explore how the contributions of individual players combine to produce a collective outcome. In a general public good game, the aggregator  $G$ —also referred to as the public good or social composition function—is represented as a Constant Elasticity of Substitution (CES) function of the players' strategies:

$$G = \left( \sum_{i=1}^n \beta_i \cdot g_i^\rho \right)^{\frac{1}{\rho}} \quad (13)$$

where  $\rho \in \mathbb{R}^*$  is the substitution parameter,<sup>6</sup>  $g_i$  is player  $i$ 's contribution to the public good and  $\beta_i$  represents the weight given to player  $i$ 's contribution:  $\beta_i > 0 \forall i$ . In a CES function, the substitution parameter  $\rho$  remains consistent across all players. We will see that together with the set of parameters  $\{\beta_i\}_{i \in I}$ , it characterises the **task type**: additive, conjunctive, disjunctive or discretionary.

Note that in the context of general public good games, the strategies consist of gifts or contributions to the public good and are usually denoted by  $g_i$ . In line with the definition of a generalised aggregative game (Definition 3.1), we see that in this case,  $G = H(\sum_{i=1}^n h_i(g_i))$  with  $H(z) = z^{1/\rho}$  and  $h_i(g_i) = \beta_i \cdot g_i^\rho$ . Furthermore,  $g_i \in \mathbb{R}_0^+$ .

This section introduces a teamwork model based on general public good games: teamwork games. The fundamental idea is that the output of teamwork can be seen as a public good to which team members have the option to contribute. When individuals

<sup>5</sup> A good is a commodity or service that can be utilised to satisfy human wants, and that has exchange value.

<sup>6</sup> In this paper, we denote by  $\mathbb{R}^+$  the set of positive real numbers,  $\mathbb{R}_0^+$  the set of non-negative real numbers including zero, and  $\mathbb{R}^*$  the set of all real numbers except zero.

work together, they subliminally face a dilemma: to spend their time on activities that yield individual benefits (e.g., leisure) while delegating the task to others, or to strive for the common good (at the expense of their time and energy). The former choice carries the risk that everyone opts for the same strategy, leading to the task not being completed, with subsequent penalties from a manager. The latter entails the risk of being the sole bearer of the collective enterprise on one's shoulders (sucker's payoff). Players' agency is manifested through the time they choose to allocate to the task. Therefore, in a teamwork game, a player  $i$ 's strategies or actions are now abstract and dimensionless, representing the percentage of a turn that the player allocates to the task. The contribution or gift  $g_i$  is calculated based on this action, but it is not a strategy per se, but a consequence of it. Thus, adjustments to the standard notation presented in Subsection 2.1 are necessary to suit the specifics of our model. In the remainder of this section, we carry out those adjustments and elucidate the pivotal elements of our model.

#### 4.1. Model

We adopt productivity as our metric to evaluate teamwork, drawing from Steiner's theoretical foundation [35]. Given a unitary task, we assume an objective and quantitative measure of the team's productivity in terms of work units, denoted by  $G$  and given by Eq. (13).

**Definition 4.1 (Unitary Team Task).** Tasks can be completed through a sequence of rounds or turns. A unitary team task  $\mathcal{T}$  is characterised by a scalar  $G^{\mathcal{T}} \in \mathbb{R}_0^+$  representing the team productivity (in work units) needed to complete the task, a pair  $[\Delta, N] \in \mathbb{N}^2$  indicating **1**) the length (in time units) of each turn and **2**) the duration (in turns) of a task, and a set of parameters  $\Theta^{\mathcal{T}} = (\rho, \beta_1, \dots, \beta_n)$  (with  $\rho \in \mathbb{R}^*$  and  $\beta_i \in \mathbb{R}^+ \forall i$ ) which specify the task's typology (additive, conjunctive, disjunctive, or discretionary):

$$\mathcal{T} := (G^{\mathcal{T}}, [\Delta, N], \Theta^{\mathcal{T}})$$

Given a vector of contributions to the task  $(g_1, \dots, g_n)$  resulting in  $G$  work units as per Eq. (13), the team task will be deemed accomplished after  $t$  turns if the team's productivity units suffice to complete the task within the designated timeframe. Formally, the following conditions must be met simultaneously:

$$\begin{cases} G \geq G^{\mathcal{T}} \\ t \leq N \end{cases} \quad (14)$$

Players exhibit varying levels of proficiency in completing tasks. When two players with different ability levels dedicate the same amount of time to a task, their contributions will inevitably differ. More precisely, we define *expertise* as the derivative of productivity with respect to time:

**Definition 4.2 (Expertise Concerning a Unitary Task).** For a player  $i$ , their expertise  $p_i^{\mathcal{T}} \in [0, 1]$  regarding a unitary task  $\mathcal{T}$  quantifies the amount of work  $i$  can accomplish per unit of time allocated to that task:

$$p_i^{\mathcal{T}} = \left( \frac{dg_i}{dt} \right) \geq 0 \quad (15)$$

When a team consists of players with varying levels of expertise, we describe it as **heterogeneous**. Conversely, if all players have the same level of expertise, the team is described as **homogeneous**. We will describe a team's joint expertise by the vector of expertise levels:  $\mathbf{p} = (p_1, \dots, p_n)$ .

Each turn  $t$ , players' agency manifests through the percentage  $a_i$  of that turn they choose to devote to the task. Therefore, a player  $i$ 's strategies or actions  $\mathcal{A}_i \subseteq \mathbb{R}_0^+$  are abstract and dimensionless, representing the percentage of the total turn that a player allocates to the task. A joint action is then a vector of these strategic choices:  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}$ , where  $\mathcal{A} \equiv \prod_{i=1}^n \mathcal{A}_i$ . Consequently, a player  $i$  with expertise  $p_i^{\mathcal{T}}$  who chooses an action  $a_i$  in a turn of duration  $\Delta_t$ , makes a contribution or gift of  $g_i$  work units:

$$g_i = a_i \cdot p_i^{\mathcal{T}} \cdot \Delta_t \quad (16)$$

In a teamwork game, the aggregative public good  $G$  represents the teamwork outcome and is given by a Constant Elasticity of Substitution (CES) function just like in a general public good game. The only difference is that now the contributions are computed as in Eq. (16):

$$G = \left( \sum_i^n \beta_i \cdot g_i^\rho \right)^{\frac{1}{\rho}} = \left( \sum_i^n \beta_i \cdot (a_i \cdot p_i^{\mathcal{T}} \cdot \Delta_t)^\rho \right)^{\frac{1}{\rho}} \quad (17)$$

Setting the parameters of this function appropriately, we obtain a teamwork model incorporating the main ingredients of Steiner's theory of group productivity:

1.  $\beta_i$  represents the weight given to the contribution of team member  $i$ .

2. Together with the set of parameters  $\{\beta_i\}_{i \in I}$ , the substitution parameter  $\rho$  characterises the **task type**: additive, conjunctive, disjunctive or discretionary (Section 2). It remains consistent across all players.
3. **Additive** tasks are those where  $\rho = 1$  and  $\beta_i = 1 \forall i$ . In this case, all contributions to the public good are aggregated additively. This setting reverts to the canonical public goods model [47,37].
4. **Conjunctive** tasks correspond to  $\rho < 1$ : the lower a contribution within the set of players, the higher its influence on the team outcome. As  $\rho \rightarrow -\infty$  with  $\beta_i = 1 \forall i$ , the task becomes strongly conjunctive, where the poorest individual performance entirely determines the team productivity. Mathematically, an aggregative public good with  $\rho < 1$  is concave in  $g_i$  [36].
5. **Disjunctive** tasks are with  $\rho > 1$ : the higher a contribution within the set of players, the higher its influence on the team outcome. As  $\rho \rightarrow +\infty$  and  $\beta_1 = \beta_2 = \dots = \beta_n = 1$ , the task becomes strongly disjunctive, where the best individual performance entirely determines the team productivity. In other words, the team's result is determined by the player who generates the most work units  $g_i$ . Mathematically, an aggregative public good with  $\rho > 1$  is not concave in  $g_i$  [36].
6. Finally, a **discretionary** task corresponds to a CES function with  $\rho = 1$  and  $\beta_1, \dots, \beta_n > 0$ , with at least some  $i, j \in I \mid \beta_i \neq \beta_j$ . This means that the team outcome is a weighted sum of the individual contributions.

Additive, conjunctive, and disjunctive tasks have been far more extensively studied than discretionary ones in the literature of social psychology, both in theoretical and experimental contexts. Consequently, we will solely focus on these three typologies, leaving the study of discretionary tasks for future work.

Just as expertise measures the productivity of a player per unit of time, a player's *leisure capacity*  $p_i^L$  measures a player's efficiency in deriving personal benefit from their free time. Here, leisure is defined as any activity a player can engage in while delegating a task to others, yielding individual enjoyment or utility outside of team responsibilities. Importantly, we assume leisure can be quantified in units similar to those used for work-related tasks, denoted by  $\mathcal{L}$ , so that both work and leisure activities are comparable in terms of time allocation and outcomes.

In this framework, *leisure capacity* represents the rate at which a player can convert free time into leisure benefits. This is analogous to *expertise* (Definition 4.2), which represents the rate of work output per unit of time for a specific task. By defining both leisure capacity and expertise in comparable terms, we create a consistent basis for evaluating how players allocate their time between group-oriented and individual activities.

**Definition 4.3 (Leisure Capacity).** Leisure capacity quantifies a player's ability to efficiently obtain leisure units during their free time. Denoted as  $p_i^L \in [0, 1]$ , it is calculated as follows:

$$p_i^L = \left( \frac{d\mathcal{L}}{dt} \right)_i \geq 0 \quad (18)$$

In other words, Equation (18) defines leisure capacity as the rate at which player  $i$  can achieve leisure benefits,  $\mathcal{L}$ , when allocating free time to personal activities. This rate-based approach allows us to directly compare leisure capacity with expertise, as both are expressed as quantities per unit of time.

With this definition, we can write the private contribution  $x_i$  representing leisure units allocated to private activities by  $i$  as:

$$x_i = (1 - a_i) \cdot p_i^L \cdot \Delta_i = p_i^L \cdot \Delta_i - g_i \frac{p_i^L}{p_i^T} \quad (19)$$

Where in the second equality we have used Eq. (16). Notice how  $\Delta_i \cdot p_i^L$  plays the role of  $i$ 's exogenous income. This expression of the private contribution allows us to write the following budget constraint for teamwork games:

**Assumption 4.1 (Individual budget constraints).** Suppose  $i$ 's exogenous income is  $\Delta_i \cdot p_i^L$ . Then, player  $i$ 's budget constraint requires that

$$\Delta_i \cdot p_i^L = x_i + g_i \frac{p_i^L}{p_i^T}$$

Going forward, we denote  $\Theta^T$  as  $\Theta$  unless specified otherwise. As already stated in Section 1, in this paper we use the full array of tools in aggregative games to analyse teamwork as a one-shot game. Thus, we will concentrate on one-shot tasks:  $\mathcal{T} = (G^T, [\Delta_i, N = 1], \Theta)$ . Furthermore, for the sake of simplicity, we adopt the convention that  $p_i^L = 1$  for all  $i$  (i.e., we assume that all players are capable of fully enjoying their free time):

$$\Delta_i \equiv x_i + \frac{g_i}{p_i^T} \quad (20)$$

Solving for  $x_i$  in Eq. (20) and adding  $G$  to both sides yields what is known as player  $i$ 's full income  $M$ :

**Definition 4.4 (Full Income).** A player's full income comprises the sum of their free time and the accessible public good, encompassing all resources available to the player.

$$M \stackrel{\text{def}}{=} x_i + G \equiv \Delta_i - \frac{g_i}{p_i^\tau} + (G_{-i}^\rho + \beta_i g_i^\rho)^{\frac{1}{\rho}} \quad (21)$$

where we have used that  $G = (G_{-i}^\rho + \beta_i g_i^\rho)^{\frac{1}{\rho}}$ . Notice that  $M$  incorporates the contributions of others explicitly as a component of player  $i$ 's income endowment, which will prove to be useful in forthcoming sections.

Because teamwork games will be defined as general public good games with further restrictions, it is noteworthy to mention other two assumptions applying to general public good games:

**Assumption 4.2 (Well-behaved preferences).** In a general public good game, the (reduced) utility is a function  $\tilde{u}_i$  of  $i$ 's private good and of the public good  $G$  for all  $i$ :

$$\tilde{u}_i(x_i, G)$$

This utility is everywhere strictly increasing and strictly quasi-concave in both arguments. It is also continuously differentiable.

**Assumption 4.3 (Normality).** In a general public good game, both the private good and the public good are normal for every player (Definition 3.6).

As anticipated in Section 1, one of the contributions of this work is an expansion of [36] that explicitly incorporates diverse teamwork evaluation functions designed to assess team performance. Having presented the key elements of our model, we now formalise the concept of teamwork evaluation.

**Definition 4.5 (Evaluation Function).** Let  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in I}$  denote a general public good game, where  $\mathcal{A}_i$  represents player  $i$ 's action set, and  $G(\mathbf{a}) \in \mathcal{G}$  is the public good (team outcome) that the set of players produces according to Eq. (17). A function  $\sigma : \mathcal{G} \rightarrow [p, q] \subset \mathbb{R}_0^+$  is an evaluation function if it satisfies the following properties:

1.  $\sigma$  is monotonically strictly increasing.
2.  $\sigma$  is differentiable.
3.  $\sigma$  has a second derivative  $\sigma''$  that satisfies the following inequality for all  $G$  in the domain of  $\sigma$ :

$$\sigma'(G)^2 - \sigma''(G) \cdot \sigma(G) > 0 \quad (22)$$

The infimum  $p$  of  $\sigma$  denotes the minimum assessment that teamwork can achieve, while the supremum  $q$  represents the maximum attainable assessment.

The introduction of the evaluation function is crucial in our context because teamwork games differ from general public good games in a significant way: a player's utility depends not only on their contribution but also on how the evaluation function evaluates the team's performance. With all the aforementioned components in place, we can now define a teamwork game as follows:

**Definition 4.6 (Teamwork Game).** Consider a set of players  $I = \{1, \dots, n\}$  engaged in a general public good game  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in I}$ . Given a one-shot task  $\mathcal{T} := (G^\tau, [\Delta_i, N = 1], \Theta)$ , an evaluation function  $\sigma : \mathcal{G} \rightarrow [p, q] \subset \mathbb{R}$  (Definition 4.5), a joint expertise  $\mathbf{p}^\tau = (p_1^\tau, \dots, p_n^\tau)$ , and a vector  $\mathbf{p}^L = (p_1^L, \dots, p_n^L)$  representing the set of players' leisure capacities, we define the triplet  $(\Gamma, \mathcal{T}, \sigma)$  as a *teamwork game* if it satisfies the following conditions:

1. The game is a static (simultaneous), non-cooperative game where strategic choices are represented by the percentages  $(a_1, \dots, a_n) \in \mathbb{R}_0^+$  of the total available time  $\Delta_i$  allocated by each player to the task.
2. The strategic choices  $(a_1, \dots, a_n)$  correspond to individual contributions  $(g_1, \dots, g_n)$  through Equation (16). Thus,  $g_i \geq 0 \forall i$ .
3. The aggregator  $G : \mathcal{A} \rightarrow \mathbb{R}$  is a CES function with parameters  $\Theta = \{\rho, \beta_1, \dots, \beta_n\}$ :

$$G = \left( \sum_i^n \beta_i \cdot g_i^\rho \right)^{\frac{1}{\rho}}$$

4. For a player  $i$ , their private good  $x_i$  is given by Equation (19).
5. Player  $i$ 's reduced utility (or utility for short) is a function of their private good  $x_i$  and of the assessment  $\sigma(G)$  of the form

$$\tilde{u}_i(x_i, \sigma(G)) \quad (23)$$

where  $\tilde{u}_i(x_i, G)$  was the original utility function of the general public good game  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in I}$ .

Building upon Definitions 4.5 and 4.6, we introduce the following Lemma to ensure that Assumptions 4.2 and 4.3 about general public good games still apply to teamwork games:

**Lemma 4.1.** Consider a teamwork game  $(\Gamma, \mathcal{T}, \sigma)$  as defined in Definition 4.6. Suppose the individual private contribution  $x_i$  and the teamwork outcome  $G$  are normal and desirable goods under certain preferences represented by the utility function  $\tilde{u}_i(x_i, G)$  in the general public good game  $\Gamma$ . Then, the following assertions hold:

1. Player  $i$ 's utility function  $\tilde{u}_i(x_i, \sigma(G))$  in the teamwork game  $(\Gamma, \mathcal{T}, \sigma)$  is differentiable and quasi-concave.
2. The private contribution  $x_i$  and the teamwork outcome  $G$  remain normal goods under this utility function.

The interested reader can find proof of the above Lemma in Appendix A. Because of the mapping between  $\mathbf{a} = (a_1, \dots, a_n)$  and  $\mathbf{g} = (g_1, \dots, g_n)$ , a joint action  $\mathbf{a}$  is fully specified by the latter without ambiguity. Consequently, when computing the Nash Equilibria of a teamwork game in the rest of the paper, we will provide the best responses in terms of gifts  $g_i$  instead of percentages of turn  $a_i$  unless stated otherwise. Specifically, we will make a slight abuse of notation and assume that the reduced best-response correspondence (Eq. (5))  $\tilde{b}_i(G_{-i}) : \mathbb{R} \rightarrow \mathcal{A}_i$  and the replacement correspondence (Definition 3.3)  $r_i(G) : \mathcal{G} \rightarrow \mathcal{A}_i$  output individual contributions  $g_i$  (instead of actions  $a_i$ ). Making this mild departure from notation conventions will provide greater clarity when discussing the impact of introducing an evaluation function into the game.

#### 4.2. Case study

Imagine you are taking an algebra course as part of your undergraduate degree. Among the tasks required to pass the subject, there is a group assignment that consists of working with four classmates to collectively solve a mathematical proof. According to the professor's instructions, only one solution needs to be submitted on behalf of the entire group. Whatever grade the exercise merits, all group members will receive the same mark. The professor informs you that passing requires completing the proof. Additionally, there is an opportunity to attempt extra exercises for bonus marks. These additional exercises increase in difficulty, and if your group fails to complete the proof, you will not be able to attempt them.

Let us assume that the team's performance can be objectively measured in units of work,  $G$ , and the professor evaluates your submission using the following sigmoid function:

$$\sigma(G) = \frac{d}{1 + e^{-\gamma(G-b)}}$$

where the minimum possible grade is  $p = 0$  and the maximum is  $q = d$ . The parameter  $b$  represents the amount of work needed to pass ( $G^T$  in Definition 4.1) with a grade of  $d/2$ . In other words, these are the units of work required to complete the proof. The parameter  $\gamma$  determines the steepness of the function. Note that if the group invests more than  $G^T$  units of work, the grade will increase, but with a decreasing rate according to the sigmoid function.

Not all group members have the same skill level for solving the proof. For instance, assume that member 1 can convert their effort into effective work with 20% efficiency (expertise), i.e.,  $p_1 = 0.2$ . This means that for every hour spent on the proof, only 20% of that time contributes to making progress. Similarly, member 2 has an expertise of 50%,  $p_2 = 0.5$ , member 3 has 60%,  $p_3 = 0.6$ , member 4 has 70%,  $p_4 = 0.7$ , and you have 80%,  $p_5 = 0.8$ . In other words, for each hour you dedicate to the proof, you effectively use 80% of your time to advance the task.

In addition to this assignment, the group members have other commitments. Some of you are taking a deep learning course, which also requires a project to be submitted on the same date, while in your free time, you attend ballet classes that you do not want to miss. In general, let  $x_i$  denote the time that the  $i$ -th group member allocates to activities other than the algebra proof.

Assume that your preferences can be expressed using a utility function based on the grade  $\sigma(G)$  and leisure  $x_i$ :

$$\tilde{u}_i(x_i, \sigma(G))$$

Being rational, each group member decides how much time to invest in the task to maximise this utility function. When you meet and start working on the task, member 4 is highly motivated and, despite being less skilled than you at proofs, willingly spends an extra hour in the afternoon working on the task. Initially, everyone worked on solving the proof in parallel. However, member 4 finishes the proof and moves on to additional exercises. The rest of the group adopts a passive role while they complete two more extra exercises. At the end of the afternoon, everyone agrees that their calculations are correct, and you decide to submit their solution as the final group assignment.

The described scenario illustrates a disjunctive task, where the overall group outcome heavily relies on the contributions of the most productive team member. A teamwork game simplifies such scenarios by modelling the decision of how much time each member contributes as a one-shot decision-making process. Since this is a disjunctive task, the team performance is best captured by employing a substitution parameter  $\rho > 1$  in the public goods production function:

$$G = \left( \sum_{i=1}^n g_i^\rho \right)^{\frac{1}{\rho}}, \quad \rho > 1,$$

where  $\rho$  controls the degree to which the highest contribution dominates the group's overall output. The use of  $\rho > 1$  reflects the unequal weighting of individual inputs characteristic of disjunctive tasks, ensuring that the outcome is primarily determined by the member who contributes the most.



## 5. Equilibria in teamwork games

In this section, we examine the equilibria in teamwork games using the aggregative games toolbox. As stated in Section 2, the (reduced) utility function of player  $i$  in aggregative games,  $\tilde{u}_i$ , differs from the usual form of utility in game theory in that its arguments are just two:  $i$ 's strategy,  $a_i$ , and some aggregate  $G$  of all the players' strategy profiles. This allows us to compute  $i$ 's best response as a function (or correspondence) of the aggregate  $G$ , without needing to distinguish the individual strategic choices of their opponents.

While this aggregative approach allows computations to scale as the number of players increases, it requires using an alternative to the best-response function, known as the replacement correspondence. The replacement correspondence  $r_i(G)$  of player  $i$  gives the set of strategies that the player would choose from in an equilibrium where the aggregate of all the contributions (including  $i$ 's) is  $G$ . Specifically, if  $G = \sum_{i=1}^n g_i$  and the replacement correspondence is a function, then  $g_i = r_i(G)$  satisfies  $g_i = \tilde{b}_i(G - g_i)$ , where  $\tilde{b}_i$  is the reduced best-response function<sup>7</sup>. We will see that when the task is additive or conjunctive, a set of first-order condition equations can be applied to directly compute the replacement correspondence, and this correspondence will be a function. However, this is not the case with disjunctive tasks: the replacement correspondence will map a set of best strategies, rather than a unique strategy. Furthermore, the first-order conditions will not directly yield the replacement correspondence for disjunctive tasks, necessitating further analysis to find it. Before exploring the formalities of equilibria in teamwork games, consider the example of the algebra course group assignment introduced in Subsection 4.2. In this scenario, each group member must decide how much time to allocate to solving the proof, balancing personal commitments with the team's shared objective. In the example, member 4 takes the initiative to invest additional time completing the proof and even tackles extra exercises, despite having less expertise than some of their peers. Meanwhile, the rest of the group adopts a passive role, focusing solely on validating member 4's results. This dynamic illustrates the disjunctive nature of the task, where the performance of the most productive individual strongly influences the team's overall outcome.

As we will see in this section formally, this joint strategy represents a Nash equilibrium in the context of a disjunctive task, such as the algebra exercise.

### 5.1. Conditions for the replacement correspondence to be defined

The conditions for the replacement function (or correspondence, where applicable) to be well-defined will vary from one type of task to another. This subsection presents these conditions in an orderly manner: first for additive tasks, then for conjunctive tasks, and finally for disjunctive tasks.

#### 5.1.1. Additive tasks

When the task is additive ( $\rho = 1$ ), the teamwork outcome  $G$  is a concave function in  $g_i$  for all  $i$  [36] and our teamwork game has the same aggregate  $G$  as a canonical public goods game:

$$G = \sum_{i=1}^n \beta_i \cdot g_i \quad (24)$$

This form of the aggregate together with Lemma 4.1 makes it possible to adopt a result by Cornes and Hartley [37] into teamwork games: Under assumptions [4.1-4.3] and with a CES function representing an additive task as given by Eq. (24), the replacement correspondence of player  $i$  (see Eq. (6)) is a function  $r_i : \mathcal{G} \rightarrow \mathcal{A}_i$  with the following properties:

1. There exists a finite value,  $\bar{G}_i$ , at which  $r_i(\bar{G}_i) = \frac{\bar{G}_i}{\Delta_i \cdot p_i} \forall i$ .
2.  $r_i(G)$  is defined for all  $G \geq \bar{G}_i$ .
3.  $r_i(G)$  is continuous.
4.  $r_i(G)$  is everywhere non-increasing in  $G$  and is strictly decreasing wherever it is strictly positive.

Here,  $\bar{G}_i$  denotes player  $i$ 's *standalone value*: the level of the teamwork outcome that player  $i$  would contribute if they were the sole contributor to the task [37]. Thus,  $\bar{G}_i$  represents the maximum individual contribution of player  $i$ , setting a threshold for the teamwork outcome  $G$  from which  $r_i(G)$  becomes defined. For any  $G$  below this threshold,  $r_i$  is undefined, precluding the use of the replacement function for Nash Equilibrium calculations<sup>8</sup>.

#### 5.1.2. Conjunctive tasks

When  $\rho < 1$  and  $\rho \neq 0$ , the task is conjunctive (the lower a contribution within the set of players, the higher its influence in the team outcome) and the teamwork outcome  $G$  is a concave function in  $g_i$  for all  $i$  [36]. We can adapt a central lemma in [36] to

<sup>7</sup> In the general case where  $G = (\sum_{i=1}^n \beta_i \cdot g_i^\rho)^{\frac{1}{\rho}}$ , the corresponding equality would be  $r_i(G) = \tilde{b}_i \left( \left[ G^\rho - \beta_i \cdot g_i^\rho \right]^{\frac{1}{\rho}} \right)$ .

<sup>8</sup> As noted in the previous section, we employ a slight abuse of notation by expressing the replacement correspondence in terms of contributions  $g_i$  rather than actions  $a_i$ , in order to maintain clarity in discussing teamwork outcomes.

determine the circumstances under which a well-defined replacement function exists in a teamwork game where players are involved in a conjunctive task:

**Lemma 5.1.** *Let  $(\Gamma, \mathcal{T}, \sigma)$  be a teamwork game featuring a conjunctive task. For every player, suppose that their private contribution  $x_i$  and the team outcome  $G$  are normal and desirable goods. Then, the replacement correspondence  $r_i(G)$  is a function satisfying:*

$$\left. \frac{\partial \tilde{u}_i}{\partial g_i} \right|_{g_{-i}} \geq 0 \quad (25)$$

with equality whenever  $g_i < p_i^T \cdot \Delta_i$ . Furthermore, the replacement function is well-defined and has two potential domains, each corresponding to a regime of parameter  $\rho$ .

1. if  $0 < \rho < 1$ ,  $r_i : [\bar{G}_i, \infty) \rightarrow \mathcal{A}_i$
2. if  $\rho < 0$ ,  $r_i : [0, \bar{G}_i] \rightarrow \mathcal{A}_i$

In teamwork games featuring conjunctive tasks, the replacement correspondence used to compute Nash Equilibria becomes a function. However, the domain of this function varies based on *how conjunctive* the task is. For quasi-conjunctive tasks where the teamwork outcome depends on the set of the  $n$  lowest contributions ( $0 < \rho < 1$ ), player  $i$ 's replacement function encompasses any teamwork outcome  $G$  equal to or surpassing their standalone level. Conversely, in strongly conjunctive tasks ( $\rho < 0$ ), where the teamwork outcome is predominantly determined by the team's lowest contribution, player  $i$ 's replacement function spans the interval  $[0, \bar{G}_i]$ . The interested reader can find a proof of Lemma 5.1 in Appendix B.

### 5.1.3. Disjunctive tasks

When  $\rho > 1$ , the task is disjunctive (the higher a contribution within the set of players, the greater its influence in the team outcome) and the team outcome no longer exhibits concavity in  $g_i$ . This requires an extension of the replacement function to a correspondence [36].

As warned by [Cornes and Hartley](#), outcomes under non-concave CES functions are notably sensitive to the specific choice of the utility function. In their work, [Cornes and Hartley](#) usually concentrate on Cobb-Douglas (CD) preferences, a type of utility function commonly used in economics to describe how consumers allocate their consumption across different goods (in their case, between the private good  $x_i$  and the public good  $G$ ):

$$\tilde{u}_i^{\text{CD}}(x_i, G) = x_i^\alpha \cdot G \quad (26)$$

with  $\alpha > 0$  being the same for all players  $i \in \mathcal{I}$  and capturing the relative importance of private consumption compared to the public good. These preferences are particularly suitable for articulating the tension between a player's desire for private good  $x_i$  and their desire for the availability of public good  $G$ : to maximise the utility  $\tilde{u}_i^{\text{CD}}(x_i, G)$ , a balance between  $x_i$  and  $G$  is required. An increase in one of the components will only significantly increase the utility if the other component is also sufficient. A high value of  $\alpha$  indicates that the private good is of relatively greater importance. In contrast, a low value indicates that the public good is more critical to the player's utility.

As expressed in Equation (23), teamwork games differ from general public good games in that now a player's utility hinges on their private contribution  $x_i$  and on how the evaluation function appraises the team's performance,  $\tilde{u}_i(x_i, \sigma(G))$ . In this section, we narrow our analysis and propose the following utility function, inspired by Cobb-Douglas preferences:

$$\tilde{u}_i(x_i, \sigma(G)) = x_i^\alpha \cdot \sigma(G) \quad (27)$$

This utility function bears a close resemblance to the Cobb-Douglas function but incorporates the teamwork assessment. It's worth noting that  $\alpha > 0$  remains uniform across all players, representing the relative significance of leisure to player  $i$  compared to the received evaluation. Conversely,  $\sigma(G)$  embodies the evaluation function, as outlined in Definition 4.5. Decoupling the teamwork outcome  $G$  as

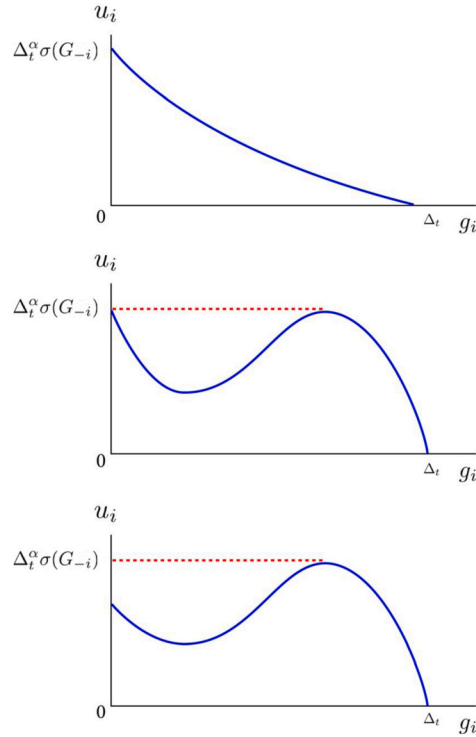
$$G = (\beta_i g_i^\rho + G_{-i}^\rho)^{\frac{1}{\rho}}$$

we can write the best response set for player  $i$  in a teamwork game with preferences given by Eq. (27):

$$B_i(G_{-i}) = \operatorname{argmax}_{g_i \in [0, \Delta_i p_i^T]} \left\{ x_i^\alpha \cdot \sigma \left( (\beta_i g_i^\rho + G_{-i}^\rho)^{\frac{1}{\rho}} \right) \right\}$$

with  $G_{-i} = \left( \sum_{j \neq i} \beta_j g_j^\rho \right)^{\frac{1}{\rho}}$ , and  $x_i$  given by Eq. (19).

Once again, we adapt a lemma by Cornes and Hartley [36] about general public good games to teamwork games to show the existence of three potential behavioural regimes within the set of best responses:



**Fig. 1.** Player  $i$ 's utility  $\tilde{u}_i(g_i, \sigma(G_{-i}))$ , across three regime values of  $G_{-i}$ . Upper panel:  $G_{-i} > G_{-i}^*$ . Middle panel:  $G_{-i} = G_{-i}^*$ . Bottom panel:  $G_{-i} < G_{-i}^*$ . Adapted from [36].

**Lemma 5.2.** Let  $(\Gamma, \mathcal{T}, \sigma)$  be a teamwork game featuring a disjunctive task ( $\rho > 1$ ):

$$G = \left( \sum_i^n \beta_i \cdot g_i^\rho \right)^{\frac{1}{\rho}} \text{ with } \rho > 1$$

Suppose that player  $i$  has preferences as described in Equation (27), where  $\sigma(G)$  denotes an evaluation function (Definition 4.5). Then, there exists a threshold value  $G_{-i}^* > 0$  and a positive real-valued function  $\tilde{b}_i$  on  $[0, G_{-i}^*]$  such that:

$$B_i(G_{-i}) = \begin{cases} \{0\} & \text{if } G_{-i} > G_{-i}^* \\ \{0, \tilde{b}_i(G_{-i})\} & \text{if } G_{-i} = G_{-i}^* \\ \{\tilde{b}_i(G_{-i})\} & \text{if } G_{-i} < G_{-i}^* \end{cases} \quad (28)$$

The interested reader will find a proof of this lemma in Appendix D.

Eq. (28) shows that there are three possible behaviour regimes within the set of best responses:

1. **No-contribution regime:** when the provision of the teamwork outcome by others surpasses a certain threshold (i.e. exceeds  $G_{-i}^*$ ), player  $i$ 's best response is not to contribute.
2. **Indifference regime:** if  $G_{-i} = G_{-i}^*$ , player  $i$  becomes indifferent between contributing or abstaining.
3. **Contribution regime:** when the teamwork outcome is suboptimally provided relative to player  $i$ 's preferences,  $i$  will contribute according to their best response.

We can visually represent player  $i$ 's utility across these scenarios based on their contribution while holding the provision of the teamwork outcome by others constant (see Fig. 1). When  $G_{-i} > G_{-i}^*$  (upper panel), maximum utility occurs at  $\tilde{u}_i(0, G_{-i}) = (\Delta_i \cdot p_i^L)^\alpha \cdot G_{-i}$ , where  $\Delta_i \cdot p_i^L$  is  $i$ 's exogenous income and  $p_i^L = 1$  for all  $i$ . At the threshold  $G_{-i}^*$ , utility maximisation is achieved at two points:  $g_i = 0$  and  $g_i = \tilde{b}_i(G_{-i}^*)$  (middle panel). Finally, when  $G_{-i} < G_{-i}^*$  (bottom panel), the highest utility is reached for some  $\tilde{b}_i(G_{-i}) > 0$ .

**Table 2**

Numerically computed Nash equilibria for a heterogeneous team with  $\mathbf{p} = (0.3, 0.8)$  across nine teamwork games of varying difficulty levels: easy, medium, and hard, for additive, conjunctive, and disjunctive tasks.

	Easy Evaluation	Medium Evaluation	Hard Evaluation
Additive ( $\rho = 1$ )	(0%, 45%)	(5%, 64%)	(33%, 75%)
Conjunctive ( $\rho = -500$ )	(56.756%, 21.351%)	(60.123%, 22.620%)	(60.198%, 22.649%)
Disjunctive ( $\rho = +500$ )	(0%, 45.1%)	(0%, 65.9%)	(0%, 82.2%)

## 5.2. An example

Before delving into the mathematical formulation of replacement correspondences in teamwork games, we showcase the types of equilibria obtained. We use Eq. (27) to model the players' preferences:

$$\tilde{u}_i(x_i, \sigma(G)) = x_i^\alpha \cdot \sigma(G)$$

For illustrative purposes, let's assume that enjoying leisure is quadratically more important to the players than receiving a good evaluation, with  $\alpha = 2$  for all players. Furthermore, we take the evaluation function (Definition 4.5) to be a logistic function:

$$\sigma(G) = \frac{d}{1 + e^{-\gamma(G-b)}}$$

The logistic function takes the teamwork outcome  $G$  as input and returns an evaluation within the interval  $[0, d] \subset \mathbb{R}$ . The parameter  $\gamma$  gives a sense of how many units of teamwork outcome are necessary to raise the assessment by one point, and in our case is set to  $\gamma = 2$ . When  $G = b$ , the teamwork outcome receives an evaluation of  $\frac{d}{2}$ , making  $b$  the passing point. Table 2 displays the Nash equilibria, calculated numerically, for a heterogeneous team with joint expertise  $\mathbf{p} = (0.3, 0.8)$  across nine teamwork games. These games vary in difficulty, covering easy, medium, and hard levels for additive, conjunctive, and disjunctive tasks.

A Nash equilibrium is represented by a vector of strategies  $\mathbf{a}$ , where  $a_1$  is the strategy of the agent with expertise  $p_1^T$  and  $a_2$  is the strategy of the agent with expertise  $p_2^T$ . As detailed in Section 4, these strategies indicate the percentage of time each player dedicates to the teamwork outcome in equilibrium.

In an additive task, the teamwork outcome results from the sum of individual efforts, with each member's contribution adding to the collective outcome. In a conjunctive task, the outcome is determined by the contribution of the team member who performs the worst. In a disjunctive task, the outcome is determined by the contribution of the member who performs the best. The table shows that increasing the level of difficulty leads to greater time contributions from the players. However, the type of task also affects the equilibria. Although the impact of task type on contributions will be detailed in Section 7, the general trend arises in this table:

- In additive tasks (top row), both members' contributions are equally important, and both increase as the evaluation becomes more difficult.
- In strongly conjunctive tasks ( $\rho \approx -\infty$ ) (middle row), only the smallest contribution determines the outcome. Even if a player wants to contribute generously to the teamwork outcome, their efforts will be wasted if someone else contributes less. Therefore, the rational strategy for players is to match the smallest contribution of the others. Indeed, computing  $a_i \cdot p_i$ , one can check that both players in the table are contributing approximately<sup>9</sup> the same.
- In strongly disjunctive tasks ( $\rho \approx \infty$ ) (bottom row), only the best contribution is added to the outcome. In this case, the teamwork game may have multiple equilibria, with only one player contributing while others free-ride. In our case, the player with more expertise is the only one who contributes in all scenarios, although this is not always the case, and in general there is more than one potential contributor.

## 5.3. Replacement correspondences and Nash equilibria in teamwork games

Once the conditions for  $r_i$  to be well-defined have been clarified, we can compute its form. We remind the reader that the output of  $r_i$  is the same as  $\hat{b}_i$ 's (i.e.  $i$ 's best possible strategy), what changes is the function's input ( $G$  vs  $G_{-i}$ ). Having the form of the replacement correspondence enables us to derive the aggregate replacement correspondence (see Definition 3.4). According to Proposition 3.1, this allows us to obtain the Nash equilibria of the game through the following procedure:

1. Compute  $R(G)$  as in Eq. (8).
2. Obtain the equilibrium aggregate  $\hat{G}$  as the fixed points of  $R(G)$ :  $\hat{G} \in R(\hat{G})$
3. By Proposition 3.1, any fixed point  $\hat{G}$  determines a set  $\{\hat{\mathbf{a}}\}$  of (possibly unique) pure strategy Nash equilibria, which we compute by applying the set of replacement correspondences  $\{r_i\}_{i \in I}$  over  $\hat{G}$ :

$$\{\hat{\mathbf{a}}\} = \{(\hat{a}_1, \dots, \hat{a}_n) \in r_1(\hat{G}) \times r_2(\hat{G}) \times \dots \times r_n(\hat{G})\} \quad (29)$$

<sup>9</sup> We attribute the slight discrepancies in contributions to the use of numerical methods in equilibrium calculations, which may introduce small errors.

This section will now detail the construction of the replacement correspondence for each task type. As outlined previously, the resulting expression can then be directly employed to derive the NE of the game. Whether  $r_i$  is a function or a correspondence will depend on the task type (i.e. the value of  $\rho$ ). In what follows, we will take the (reduced) utility function  $\tilde{u}_i$  to be that in Eq. (27):

$$\tilde{u}_i(x_i, \sigma(G)) = x_i^\alpha \cdot \sigma(G)$$

and we will obtain the expression of the function/correspondence  $r_i$  for additive, conjunctive and disjunctive tasks.

### 5.3.1. Additive tasks

When tasks are additive ( $\rho = 1$ ), we saw in Section 5.1.1 that if the teamwork's outcome  $G$  is greater than or equal to every player's standalone value  $\bar{G}_i$ , then the replacement correspondence is a function. In those circumstances, the existence and uniqueness of an NE are guaranteed (see [37] for a thorough justification). Furthermore, the teamwork outcome  $G$  is a concave function in  $g_i$  for all  $i$  [36]. This implies that this equilibrium can be computed by working through the standard first-order conditions of utility maximisation (Eq. (25)).

Substituting  $\rho = 1$  in Eq. (13), it is easy to check that when an evaluation function over the teamwork outcome  $G$  is employed and player  $i$ 's utility is governed by Equation (27), the first-order conditions in Eq. (25) yield the following replacement function:

$$r_i(G) = \max \left\{ 0, p_i^\tau \Delta_i - \frac{\alpha}{\beta_i} \cdot \frac{\sigma(G)}{\sigma'(G)} \right\} \quad (30)$$

Looking at the non-zero term of Eq. (30), the minuend represents the potential productivity that player  $i$  could achieve if they dedicated an entire turn  $\Delta_i$  to contributing to the task. The subtrahend quantifies the extent to which this is not possible: higher values render the subtrahend more negative, indicating a greater impediment to the agent to contribute.

Thus, when  $p_i^\tau \Delta_i$  is less than  $\frac{\alpha}{\beta_i} \cdot \frac{\sigma(G)}{\sigma'(G)}$ , the result of  $r_i(G)$  is zero and player  $i$  chooses not to contribute. That is, when the expertise is sufficiently low, the best a player can do is free-ride on the efforts of others with greater expertise than themselves. Conversely, if  $p_i^\tau \Delta_i$  is greater than  $\frac{\alpha}{\beta_i} \cdot \frac{\sigma(G)}{\sigma'(G)}$ , then  $r_i(G) = p_i^\tau \Delta_i - \frac{\alpha}{\beta_i} \cdot \frac{\sigma(G)}{\sigma'(G)}$ , indicating that player  $i$  contributes a positive amount determined by this difference. Holding other parameters constant, as  $p_i^\tau$  increases, it becomes more likely for player  $i$  to contribute more, as  $p_i^\tau \Delta_i$  becomes larger in comparison to the negative term in the equation. The same applies to the behaviour of  $r_i(G)$  as  $\Delta_i$  increases. In short, the higher a player's expertise, the more likely they are to contribute. Another way to encourage a player to contribute is by providing them with more time  $\Delta_i$  (longer turns).

As for  $\alpha$ , it captures the relative importance that players give to their private time compared to the teamwork outcome. For the case  $r_i(G) \neq 0$ , it becomes evident that, for a fixed quantity of teamwork outcome  $G$  and a fixed weight  $\beta_i$ , the optimal contribution  $g_i$  diminishes as  $\alpha > 0$  increases. This is an intuitive outcome, as higher values of  $\alpha$  signify a greater emphasis on leisure in the preferences of player  $i$ , thereby discouraging contributions.

Conversely, as  $\beta_i$  increases (with all other factors held constant), the absolute value of the negative term diminishes. Put differently, the greater the importance or weight  $\beta_i$  assigned to a player's contributions, the more the player will contribute, assuming other factors remain constant.

Regarding the teamwork evaluation function, higher values of the fraction  $\frac{\sigma(G)}{\sigma'(G)}$  render the subtrahend in Eq. (30) more negative, indicating a greater impediment to the agent's contribution. This ratio can increase either due to a rise in the numerator or a decrease in the denominator. For simplicity, let us consider for a moment that either the numerator or the denominator changes while the other remains constant.

If the numerator increases at a point  $G_0$  while  $\sigma'(G)$  remains constant, it means that the evaluation has become more lenient, or in other words, the passing threshold has been lowered, making it easier to pass. Perhaps counterintuitively, this lowering of the passing threshold inhibits player contributions rather than incentivising them, as the same utility is obtained with less effort than before.

If we fix the value of the assessment at some teamwork outcome value,  $\sigma(G_0)$ , but decrease the slope of the evaluation function  $\sigma'(G)$ , then more teamwork outcome units  $\Delta G$  will be needed to increase the evaluation from  $\sigma(G_0)$  to  $\sigma(G_0) + 1$ . In this situation, a player would prefer an evaluation function with a steeper slope, as an increase in their contributions would result in greater utility. In other words, decreasing the slope of the evaluation at a point on the curve  $\sigma(G_0)$  inhibits contributions.

In both cases, the fraction  $\frac{\sigma(G)}{\sigma'(G)}$  increases. In conclusion: an easier pass or a less steep evaluation function inhibits players from contributing more. And vice-versa: an optimal contribution  $g_i$  will increase as the fraction decreases, as the opportunity cost in terms of leisure time (i.e., it is easier to improve the utility with the same amount of effort).

### 5.3.2. Conjunctive tasks

When  $\rho < 1$  and  $\rho \neq 0$ , the task is conjunctive, and the teamwork outcome  $G$  is a concave function in  $g_i$  for all  $i$  (i.e. similar to the additive case, equilibria can be computed by working through the standard first-order conditions of utility maximisation). Lemma 5.1 states that in this scenario, the replacement correspondence is a function (with two possible domains depending on the value of  $\rho$ ) satisfying the first-order conditions (FOCs) in Eq. (25). In an analysis we do not replicate here due to space constraints, Cornes and Hartley employ the share function (Definition 3.5) to demonstrate the existence and uniqueness of an equilibrium in general public good games with concave  $G$ . We can extend this finding to our context:



**Proposition 5.1.** *Let  $(\Gamma, \mathcal{T}, \sigma)$  be a teamwork game featuring a conjunctive task. Under the conditions of Lemma 5.1, the game has a unique equilibrium whenever the utility function  $\tilde{u}_i(x_i, \sigma(G))$  is such that the corresponding indifference map is asymptotic to the axes.*

The interested reader can find a proof of Proposition 5.1 in Appendix B. Applying the FOCs in Eq. (25) and using the utility function in Eq. (27), best-responses can be computed with the following equation:

$$\left. \frac{\partial \tilde{u}_i}{\partial g_i} \right|_{g_{-i}} = \frac{\partial \tilde{u}_i}{\partial x_i} \frac{\partial G}{\partial g_i} \left\{ \frac{x_i}{\alpha} \frac{\sigma'(G)}{\sigma(G)} - \left( \frac{\partial G}{\partial g_i} \right)^{-1} \right\} \geq 0 \quad (31)$$

with equality for  $g_i < p_i^\mathcal{T} \cdot \Delta_i$ . Manipulating, we obtain the following expression for the replacement function in teamwork games with conjunctive tasks:

$$\frac{\sigma(G)}{\sigma'(G)} \cdot G^{\rho-1} = \left( \Delta_i - \frac{r_i(G)}{p_i^\mathcal{T}} \right) \cdot \frac{\beta_i p_i^\mathcal{T}}{\alpha} \cdot r_i(G)^{\rho-1} \quad (32)$$

Appendix C examines how  $r_i(G)$  changes as the parameters in this equation increase or decrease. To give a flavour of the resulting equilibria, let's study the case of strongly conjunctive tasks ( $\rho \rightarrow -\infty$ ). For simplicity, let's consider a homogeneous team with  $p_i^\mathcal{T} = 1$  for all  $i$ , and assume that all team members have equal importance:  $\beta_i = 1$  for all  $i$ . In the limit  $\rho = -\infty$ , the teamwork outcome is solely determined by the contribution of the team member with the lowest performance:

$$G = \min_{i \in I} g_i$$

By Proposition 5.1, the game has a unique equilibrium for  $\rho \rightarrow -\infty$ . Let  $\hat{G}(\rho)$  be the teamwork outcome at equilibrium. Since all players are identical, their contributions are the same due to symmetry and are given by

$$\hat{g} = \left( \frac{1}{n} \right)^{1/\rho} \hat{G}(\rho)$$

Setting the bracketed term to zero in Eq. (31) and using  $\hat{g} = \left( \frac{1}{n} \right)^{1/\rho} \hat{G}(\rho)$ , we obtain the equation that the best-responses must satisfy in the case of symmetric players:

$$\frac{\Delta_i - n^{-1/\rho} \hat{G}(\rho)}{\alpha} \cdot \frac{\sigma'(G)}{\sigma(G)} = n^{1-\frac{1}{\rho}}$$

Taking the limit  $\rho \rightarrow -\infty$ , we get the expression for  $\hat{G}(-\infty)$ :

$$\hat{G}(-\infty) = \Delta_i - n\alpha \frac{\sigma(G)}{\sigma'(G)} \quad (33)$$

Thus, in the case of tasks where the teamwork outcome is solely determined by the lowest individual contribution (and players are identical), all team members contribute the same, and the teamwork outcome becomes a function of the number of players. Specifically,  $\hat{G}(-\infty)$  decreases to zero as  $n \rightarrow \infty$ . This implies that when a task is strongly conjunctive and there is a homogeneous team, increasing the team size results in a loss of productivity at the Nash equilibrium. This observation aligns with Steiner [35]'s assertion on the effect of team size on conjunctive tasks: "when tasks are truly conjunctive (*i.e.*, when success depends on the poorest member of the group), potential productivity decreases as the group is enlarged".

### 5.3.3. Disjunctive tasks

In this section, we will explore the dynamics of teamwork games involving disjunctive tasks, and we will see that such games can result in multiple equilibria. In these equilibria, some players contribute while others defect. The analysis required to reach this conclusion is more complex than the analyses from Sections 5.3.1 and 5.3.2. This complexity arises because disjunctive tasks operate within a parameter regime of  $\rho > 1$ , resulting in teamwork outcomes that are non-concave functions of the individual contributions  $g_i$ .

Due to the non-concavity of  $G$  in  $g_i$ , equilibria cannot be determined using the standard first-order conditions for utility maximization (Eq. (25)). Instead, a more in-depth analysis is necessary to identify these equilibria. However, it is still true that for any player  $i$ , a local maximum of their utility function is reached when  $\frac{d\tilde{u}_i}{dg_i} = 0$ . Upon developing and simplifying this equation, we obtain the same expression as Eq. (32) (but now for  $\rho > 1$ ):

$$\frac{\sigma(G)}{\sigma'(G)} \cdot G^{\rho-1} = \left( \Delta_i - \frac{g_i}{p_i^\mathcal{T}} \right) \cdot \frac{\beta_i p_i^\mathcal{T}}{\alpha} \cdot g_i^{\rho-1} \quad (34)$$

Or equivalently:

$$G = \left( \frac{\beta_i p_i^\mathcal{T}}{\alpha} \cdot \left( \Delta_i - \frac{g_i}{p_i^\mathcal{T}} \right) \cdot \frac{\sigma'(G)}{\sigma(G)} \right)^{\frac{1}{\rho-1}} \cdot g_i \quad (35)$$

A teamwork contribution  $g_i$  satisfying this local maximum equation will also be a global maximum if the payoff at it is greater than or equal to the payoff at  $g_i = 0$ :

$$\left(\Delta_i - \frac{g_i}{p_i^T}\right)^\alpha \cdot \sigma(G) \geq \Delta_i^\alpha \cdot \sigma(G_{-i}) \quad (36)$$

Notice that by definition, the values of  $G_{-i}^*$  from Lemma 5.2 and the corresponding critical point  $g_i^*$  are obtained when Eqs. (34) and (36) are satisfied with equality. Because in this case, the set of best-responses can have more than one element according to Lemma 5.2, a replacement **correspondence** has to be defined in turn:

$$r_i(G) = \left\{ g_i : g_i \in B_i \left( \left[ G^\rho - \beta_i g_i^\rho \right]^{\frac{1}{\rho}} \right) \right\} \quad (37)$$

Hence, in a disjunctive task, player  $i$  might choose from multiple possible strategies in an equilibrium resulting in total teamwork outcome  $G$ . From Lemma 5.2, we know that  $g_i \in r_i(G)$  if and only if either:

$$\begin{cases} g_i = 0 \text{ and } G_{-i} \geq G_{-i}^*, \text{ or} \\ g_i > 0 \text{ and } \begin{cases} \frac{\sigma(G)}{\sigma'(G)} \cdot G^{\rho-1} = \left(\Delta_i - \frac{g_i}{p_i^T}\right) \cdot \frac{\beta_i p_i^T}{\alpha} \cdot g_i^{\rho-1} \\ \beta_i^{1/\rho} g_i \leq G \leq \left( \sigma^{-1} \left( \sigma(G) \left(1 - \frac{g_i}{p_i^T \Delta_i}\right)^\alpha \right)^\rho + \beta_i g_i^\rho \right)^{\frac{1}{\rho}} \end{cases} \end{cases} \quad (38)$$

Where for the second inequality of the case  $g_i > 0$ , we have used the global maximum condition from Eq. (36). The expressions in (38) tell us that the replacement correspondence  $r_i(G)$  in our case has two components:

- A component extending along the  $X$ -axis to the right of  $G_{-i}^*$ .
- A positive component, which is the reflection across the  $45^\circ$  line of the segment of the teamwork outcome  $G$  that simultaneously fulfils the local and global maximum conditions (Eqs. (35) and (36) respectively).

To understand the geometry of the positive component of  $r_i$ , notice that the right-hand side of Equation (34) vanishes when  $g_i$  is either equal to 0 or to  $p_i^T \cdot \Delta_i$ , and it has a peak at  $g_i = \Delta_i p_i^T \left(1 - \frac{1}{\rho}\right)$ . Additionally, it's worth noting that:

1. The boundary line of the left-hand inequality in Eq. (38), crosses the graph of Eq. (35) at the point corresponding to the standalone value of player  $i$  (the level of the teamwork outcome that player  $i$  would contribute were they the sole contributor,  $\bar{G}_i$ ):

$$(\bar{g}_i, \bar{G}_i) = \left( \Delta_i \cdot p_i^T - \frac{\sigma(\bar{G}_i)}{\sigma'(\bar{G}_i)} \frac{\alpha}{\beta_i^{1/\rho}}, \beta_i^{1/\rho} \cdot \Delta_i \cdot p_i^T - \frac{\sigma(\bar{G}_i)}{\sigma'(\bar{G}_i)} \alpha \right)$$

2. By definition, the boundary line of the right-hand inequality in Eq. (38) crosses the graph of Eq. (35) to the left of  $(\bar{g}_i, \bar{G}_i)$ , at the critical points  $(g_i^*, G_i^*)$ , where

$$G_i^* = (G_{-i}^* + \beta_i g_i^{*\rho})^{1/\rho}$$

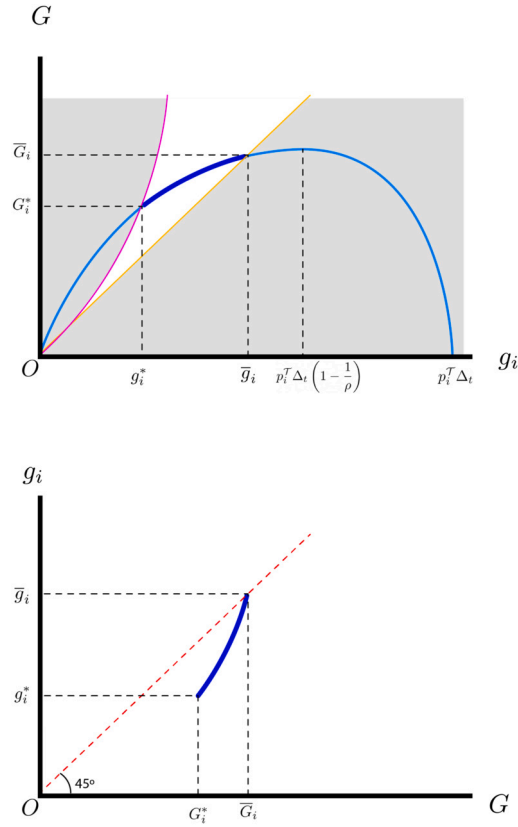
The top panel in Fig. 2 illustrates  $G$  from equation (35) and how the different boundary lines mentioned intersect with its graph. The bottom panel represents the positive component of  $r_i$  as the reflection across the  $45^\circ$  line of the segment of  $G$  that fulfils the conditions for  $g_i > 0$  in Eq. (38).

If both  $(g_i^*, G_i^*)$  and  $(\bar{g}_i, \bar{G}_i)$  are located to the left of (or at) the maximum of  $G(g_i)$ , both points will lie within the increasing portion of this function. Examining the  $45^\circ$  reflection of that function in the bottom panel of Fig. 2, this indicates that the positive component of the replacement correspondence is single-valued on its domain:  $[G_i^*, \bar{G}_i]$ . To ensure this condition is met, it is easy to check that it must satisfy:

$$\frac{\rho}{\Delta_i p_i^T} \geq \frac{\sigma'(\bar{G}_i)}{\sigma(\bar{G}_i)} \cdot \frac{\beta_i^{1/\rho}}{\alpha}$$

As a result, we get the following corollary:

**Corollary 5.1.** *In a teamwork game involving a disjunctive task, if  $\frac{\rho}{\Delta_i p_i^T} \geq \frac{\sigma'(\bar{G}_i)}{\sigma(\bar{G}_i)} \cdot \frac{\beta_i^{1/\rho}}{\alpha}$ , then the positive component of the replacement correspondence is single-valued on its domain:  $[G_i^*, \bar{G}_i]$ .*



**Fig. 2.** Top panel: The teamwork outcome  $G$ , as given by Eq. (35), is depicted in two different shades of blue and has a bell shape. The boundary line of the left-hand inequality in Eq. (38), shown in fuchsia, intersects this graph at  $(g_i^*, G_i^*)$ . The boundary line of the right-hand inequality in Eq. (38), shown in light orange, is to the right of the fuchsia line and intersects the graph at  $(\bar{g}_i, \bar{G}_i)$ . The bold blue segment within the non-shaded area indicates the specific region of  $G$  that satisfies both the local and global maximum conditions simultaneously. Bottom panel: Reflection of the teamwork outcome  $G$  across the  $45^\circ$  line, showing only the region where  $G$  meets the local and global maximum conditions. Figures adapted from [36]. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

In essence, this corollary tells us that although a player in a teamwork game involving a disjunctive task does not have a replacement function but rather a replacement correspondence as per Eq. (37), when examining the two possible components of the replacement correspondence individually, they are single-valued. One component outputs zero, and the other is a function with a domain of  $[G_i^*, \bar{G}_i]$ . As a consequence, teamwork games involving disjunctive tasks may have multiple equilibria. In each of these equilibria, some players will contribute ( $g_i > 0$ ), while all others will defect ( $g_i = 0$ ). Players who positively contribute to teamwork are called the “active” players in that equilibrium, and they form an **equilibrium set**  $\mathcal{J} \subseteq \mathcal{I}$ .

By Definition 3.5, whoever the active players are, their share correspondences will sum up to 1 in equilibrium:

$$\sum_{j \in \mathcal{J}} s_j(\hat{G}) = 1 \quad (39)$$

Whereas  $\hat{g}_i = 0 \forall i \notin \mathcal{J}$ . If we can demonstrate that the share correspondences of the active players are indeed functions and that they are increasing, then the sum above would equal 1 for each set of active players  $\mathcal{J}$  in a unique configuration of contributions. In other words, if we show that  $s_j(G)$  are increasing functions for the active players, then there will be a unique equilibrium for each set of active players. The following proposition guarantees that these conditions are met:

**Proposition 5.2.** If  $\frac{\rho}{\Delta_i p_i^\rho} \geq \frac{\sigma'(\bar{G}_i) \beta_i^{1/\rho}}{\sigma(\bar{G}_i) \alpha}$ , the graph of  $s_i(G)$  in a teamwork game featuring a disjunctive task is the disjoint union of two sets:  $\{(G, 0) : G \geq G_{-i}^*\}$  and  $\{(G, s_i(G)) : G_i^* \leq G \leq \bar{G}_i\}$ . Plus,  $s_i$  is continuous and strictly increasing, and satisfies

$$\begin{cases} s_i(G_i^*) = \beta_i \frac{g_i^{\rho}}{G_i^{*\rho}} > 0, \text{ and} \\ s_i(\bar{G}_i) = 1 \end{cases}$$

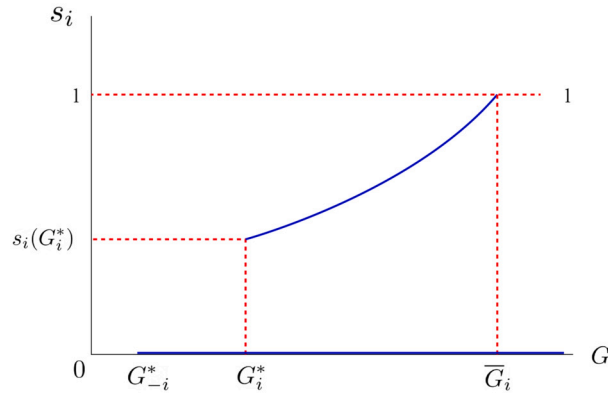


Fig. 3. The two components of player  $i$ 's share correspondence when  $\frac{\rho}{\Delta_i p_i'} \geq \frac{\sigma'(\bar{G}_i)}{\sigma(\bar{G}_i)} \frac{\beta_i^{1/\rho}}{\alpha}$ .

The interested reader can find proof of this Proposition in Appendix E. Fig. 3 depicts this result.

We can now delineate equilibria with share functions. According to Proposition 5.2 detailed above, for  $j \in \mathcal{J}$ ,  $s_j(G)$  is an increasing function in  $G$  and hence there is one equilibrium per set of active players. The level of that equilibrium is denoted by  $\hat{G}(\mathcal{J})$ , and the equilibrium is obtained through the following procedure:

1. Take  $s_i(\hat{G}(\mathcal{J})) = 0$  if  $i \notin \mathcal{J}$ .
2. Compute  $\{s_j(\hat{G}(\mathcal{J}))\}$  for  $j \in \mathcal{J}$  such that Eq. (39) is satisfied.

Therefore, characterising equilibria is equivalent to characterising equilibrium sets, with necessary conditions for an equilibrium set following from Proposition 5.2. That is, in an equilibrium set  $\mathcal{J}$ , the following conditions are fulfilled:

$$G_j^* \leq \hat{G}(\mathcal{J}) \leq \bar{G}_i \quad \forall j \in \mathcal{J}$$

$$G_{-i}^* \leq \hat{G}(\mathcal{J}) \quad \forall i \notin \mathcal{J}$$

We can use these conditions to identify equilibrium sets systematically:

1. Evaluate each possible subset  $\mathcal{J} \subseteq \mathcal{I}$ .
2. Within each subset, we select the critical value  $G_j^*$  of the player with the lowest tolerance for undersupplying the teamwork outcome.
3. Simultaneously, we choose the critical value  $G_{-i}^*$  of the player outside the subset who also exhibits the least tolerance for the teamwork outcome being undersupplied by others.
4. The maximum of these two values yields  $G^*(\mathcal{J})$ .
5. For  $\mathcal{J}$  to qualify as an active set, it must consist of players whose individual standalone values  $\bar{G}_j$  all exceed or equal this critical value  $G^*(\mathcal{J})$ , while also satisfying

$$\sum_{j \in \mathcal{J}} s_j(G^*(\mathcal{J})) \leq 1$$

More formally, let  $G^*(\mathcal{J})$  be the maximum among 1) the highest critical value  $G_j^*$  of the active players and 2) the highest critical value of the inactive players,  $G_{-i}^*$ :

$$G^*(\mathcal{J}) = \max \{ \max_{j \in \mathcal{J}} G_j^*, \max_{i \notin \mathcal{J}} G_{-i}^* \}$$

With this definition, we can adopt Cornes and Hartley's characterisation of equilibrium sets in non-concave public goods<sup>10</sup> for teamwork games with disjunctive tasks:

**Proposition 5.3.** Suppose that  $\frac{\rho}{\Delta_i p_i'} \geq \frac{\sigma'(\bar{G}_i)}{\sigma(\bar{G}_i)} \frac{\beta_i^{1/\rho}}{\alpha} \quad \forall i \in \mathcal{I}$  in a teamwork game featuring a disjunctive task. A non-empty set  $\mathcal{J} \subseteq \mathcal{I}$  is an equilibrium set if and only if it satisfies:

<sup>10</sup> See [36] for further clarification.

**Table 3**

Replacement functions/correspondences in teamwork and general public good games.

(a) Replacement functions/correspondence for the three different task types in teamwork games.	
Teamwork Games	
Additive $\mathcal{T}(\rho = 1)$	$r_i(G) = \max \left\{ 0, p_i^T \Delta_i - \frac{\alpha}{\beta_i} \cdot \frac{\sigma(G)}{\sigma'(G)} \right\}$
Conjunctive $\mathcal{T}(\rho < 1)$	$\frac{\sigma(G)}{\sigma'(G)} \cdot G^{\rho-1} = \left( \Delta_i - \frac{r_i(G)}{p_i^T} \right) \cdot \frac{\beta_i p_i^T}{\alpha} \cdot r_i(G)^{\rho-1}$
Disjunctive $\mathcal{T}(\rho > 1)$	$\begin{cases} r_i(G) = 0 \text{ and } G_{-i} \geq G_{-i}^*, \text{ or} \\ r_i(G) > 0 \text{ and } \begin{cases} \frac{\sigma(G)}{\sigma'(G)} \cdot G^{\rho-1} = \left( \Delta_i - \frac{r_i(G)}{p_i^T} \right) \cdot \frac{\beta_i p_i^T}{\alpha} \cdot r_i(G)^{\rho-1} \\ \beta_i^{\frac{1}{\rho}} r_i(G) \leq G \leq \left( \sigma^{-1} \left( \sigma(G) \left( 1 - \frac{r_i(G)}{p_i^T \Delta_i} \right)^{\alpha} \right)^{\rho} + \beta_i r_i(G)^{\rho} \right)^{\frac{1}{\rho}} \end{cases} \end{cases}$
(b) Replacement functions/correspondence for canonical, concave and convex public goods.	
General Public Good Games	
Canonical GPG ( $\rho = 1$ )	$r_i(G) = \max \left\{ 0, w_i - \frac{\alpha}{\beta_i} G \right\}$
Concave GPG ( $\rho < 1$ )	$r_i(G)^{\rho-1} (w_i - r_i(G)) = \frac{\alpha}{\beta_i} \cdot G^{\rho}$
Convex GPG ( $\rho > 1$ )	$\begin{cases} r_i(G) = 0 \text{ and } G_{-i} \geq G_{-i}^*, \text{ or} \\ r_i(G) > 0 \text{ and } \begin{cases} r_i(G)^{\rho-1} (w_i - r_i(G)) = \frac{\alpha}{\beta_i} \cdot G^{\rho} \\ \beta_i^{\frac{1}{\rho}} r_i(G) \leq G \leq \beta_i^{\frac{1}{\rho}} r_i(G) w_i^{\alpha} \left[ w_i^{\alpha \rho} - (w_i - r_i(G))^{\alpha \rho} \right]^{\frac{-1}{\rho}} \end{cases} \end{cases}$

$$\begin{cases} G^*(\mathcal{J}) \leq \min_{j \in \mathcal{J}} \bar{G}_j \\ \sum_{j \in \mathcal{J}} s_j(G^*(\mathcal{J})) \leq 1 \end{cases}$$

It follows directly from this proposition that non-empty subsets of equilibrium sets also satisfy equilibrium conditions. Specifically, we extend Corollary 6.1 from Cornes and Hartley [36] to characterise equilibrium outcomes in teamwork games involving disjunctive tasks:

**Corollary 5.2.** *If  $\frac{\rho}{\Delta_i p_i^T} \geq \frac{\sigma'(\bar{G}_i)}{\sigma(\bar{G}_i)} \cdot \frac{\beta_i^{1/\rho}}{\alpha}$  for all  $i \in \mathcal{I}$ , then for any equilibrium set  $\mathcal{J}$ , any subset  $\mathcal{K} \subset \mathcal{J}$  also constitutes an equilibrium set, with  $\hat{G}(\mathcal{K}) > \hat{G}(\mathcal{J})$ .*

In simpler terms, for any subset  $\mathcal{K} \subset \mathcal{J}$  of active players, the teamwork outcome produced by  $\mathcal{K}$  will be greater than that of  $\mathcal{J}$ . As pointed out by Cornes and Hartley [36], this suggests that free-riders benefit more when there are fewer contributors since it allows them to enjoy a larger portion of the teamwork outcome. In scenarios with fewer active players, each contributor has to take on a larger share of the workload. This increased individual contribution can lead to a more significant overall impact on the public good being produced. As a result, those who do not contribute—free riders—can gain from the increased output without sharing the effort with many others. Essentially, by reducing the number of contributors in disjunctive tasks, free-riders can maximize their benefits.

#### 5.4. Equilibria in teamwork games vs equilibria in general public good games

Before delving into the details of the multi-agent model we propose for learning equilibrium in teamwork games, let's briefly compare our results with those of Cornes and Hartley [36]. Table 3 summarizes the replacement functions (or correspondences) obtained in Subsection 5.3 for teamwork games, alongside the corresponding functions for general public good (GPG) games as studied by Cornes and Hartley [36].

In a GPG game,  $w_i$  represents player  $i$ 's exogenous income (as opposed to  $p_i^T \Delta_i$  in teamwork games). Additionally, Cornes and Hartley do not consider heterogeneity among players (i.e., players with different levels of expertise). In fact, expertise is not explicitly modelled in their game, so we assume  $p_i^T = 1 \forall i$  in GPG games.

Examining the value of the substitution parameter  $\rho$ , we can see that GPG games involving canonical public goods correspond to teamwork games involving additive tasks in our model. Similarly, there is a mapping between GPG games with concave public goods and teamwork games involving conjunctive tasks, and between GPG games with convex public goods and teamwork games involving disjunctive tasks.

The results in Table 3 demonstrate that our findings extend those of Cornes and Hartley. By setting  $\sigma(G) = G$  and  $p_i^T = 1$ , we can reproduce the expressions for  $r_i(G)$  found in Table 3b. Notice how incorporating a non-trivial evaluation function in teamwork games influences the strategy  $g_i$  selected by player  $i$  through the denominator  $\frac{\sigma(G)}{\sigma'(G)}$ . As detailed in Section 5.1.1, this ratio represents the relationship between the evaluation function  $\sigma(G)$  and the rate at which the assessment grows as  $G$  increases. Since the rate of growth  $\sigma'(G)$  is positive by definition,  $\frac{\sigma(G)}{\sigma'(G)}$  indicates how valuable it is for players to invest their time in improving their assessment



**Table 4**

Nash equilibria for two agents with expertise  $p^T = 1$  playing three types of teamwork games and a general public good (GPG) game. The required teamwork outcome for passing is  $G = 5$ .

Additive $\mathcal{T}$ (30%, 30%)	Canonical PG (20%, 20%)	Conjunctive $\mathcal{T}$ (52%, 52%)	Concave PG (20%, 20%)
		Disjunctive $\mathcal{T}$ (56%, 0%) (0%, 56%)	Convex PG (33%, 0%) (0%, 33%)

rather than enjoying their leisure time. In other words, it provides information about the *opportunity cost* of the evaluation in terms of enjoyed leisure. Consequently, holding other parameters constant, an optimal contribution  $g_i$  will increase as this fraction decreases, as it becomes easier to improve the assessment score with the same amount of effort. Table 4 presents the Nash equilibria when two agents with expertise  $p^T = 1$  play three teamwork games: one involving an additive task, another involving a conjunctive task, and a third involving a disjunctive task.<sup>11</sup> In all these games, we take the evaluation function (Definition 4.5) to be a logistic function:

$$\sigma(G) = \frac{d}{1 + e^{-\gamma(G-b)}}$$

Furthermore, we model the players' preferences with

$$\tilde{u}_i(x_i, \sigma(G)) = x_i^\alpha \cdot \sigma(G)$$

Similar to Table 2, let's assume that enjoying leisure is quadratically more important to the agents than receiving a good evaluation, with  $\alpha = 2$  for all players. To ensure illustrative clarity, we have chosen the passing threshold  $b = 5$  and the steepness parameter  $\gamma = 2$  to clearly demonstrate the dynamics of the evaluation function. These parameter values were selected to highlight how the evaluation function shapes player strategies while remaining within a plausible range. Importantly, the qualitative behaviour of the model is robust to variations in these parameters. For comparison, we include the corresponding Nash equilibria of the corresponding general public good (GPG) games.

Just as in Table 2, a Nash equilibrium is represented by a vector of strategies  $\mathbf{a}$ , indicating the percentage of time each player dedicates to the teamwork outcome in equilibrium. Notice that in the cases shown, the introduction of the evaluation function shifts the equilibria, resulting in greater time contributions compared to public good games. Furthermore, for our chosen parameters, the equilibrium in GPG games with canonical and concave public goods is identical: in both cases, both players allocate 20% of their turn to the public good.<sup>12</sup> In a teamwork game, this symmetry breaks, and players differentiate between additive and conjunctive tasks. In all the displayed equilibria in teamwork games, the team passes ( $G \geq 5$ ). Furthermore, note that both in the teamwork game involving a disjunctive task and the GPG game involving a convex public good, there are two equilibria. In each of them, one player is the sole contributor, while the other free-rides.

## 6. Multiagent multi-armed bandit learning

In previous sections, we presented a formal model of a teamwork game and characterised the replacement correspondences for different types of tasks that teams may face. These replacement correspondences allow for the characterisation of Nash Equilibria (NE) through the application of Proposition 3.1. In this section, we propose a multiagent multi-armed bandit (MA-MAB) framework where agents' learning converges to the approximate NE of teamwork games.

The rationale behind this choice of learning framework is twofold. Firstly, this approach effectively captures the exploration-exploitation dilemma faced by decision-makers with multiple choices. Secondly, as discussed by Leslie and Collins [16], action selection methods based on smooth best responses, such as the soft-max method [68], can guide a multiagent MAB system towards strategies approximating Nash equilibria. The use of smooth best responses means that Nash equilibria (NE) are no longer fixed points in the strategy space. Instead, the algorithm converges to Nash distributions—joint strategies where each agent plays smooth best responses to rewards from interactions with others. Under certain conditions, the learning algorithm can be applied blindly to any game and, if convergence occurs, a Nash distribution will have been reached [80,16].

### 6.1. Overview

Consider an agent participating in a team who must decide how much effort to contribute. As we have seen, this decision is influenced by various factors, including the team's composition, the nature of the task, and team interaction dynamics. Conceptualising this decision-making as a teamwork game, each team member's decision-making process resembles that of a multi-armed bandit problem. Here, the "arms" represent discrete values indicating the percentage of a decision maker's total available time allocated to the task. An independent learning agent, denoted by  $i$ , only learns from its own observations, actions, and rewards, without

<sup>11</sup> The agents in this example have expertise  $p^T = 1$  to ensure the results are comparable with the equilibria of a general public good game.

<sup>12</sup> Note that this does not imply that the value of the public good in equilibrium,  $\hat{G}$ , is the same. In fact, in the GPG games shown here,  $\hat{G}_{\text{canonical}} = 4$ , while  $\hat{G}_{\text{concave}} = 2$ .

taking into account experiences from other agents [20]. This means that each learning agent only considers the effects of other agents' actions as part of the environment dynamics. The following sections present the elements of a multiagent multi-armed bandit learning framework where the agents play an underlying teamwork game.

Section 6.2 presents the elements of our multiagent framework related to the environment with which the agents interact, including the conditions under which a task is considered complete and the reward signal that agents receive for their contributions to the teamwork outcome. Section 6.3 specifies the elements that define our agents, including their action selection mechanism and the learning algorithm employed.

## 6.2. Teamwork environment

Our learning environment uses a teamwork game  $(\Gamma, \mathcal{T}, \sigma)$  as the underlying game. The agents have one turn of common length  $\Delta_t$  to complete a task  $\mathcal{T} = (G^\mathcal{T}, [\Delta_t, N = 1], \Theta)$  of typology specified by the parameters in  $\Theta = \{\rho, \beta_1, \dots, \beta_n\}$ . Because the task is one-shot, the fulfilment conditions in Eq. (14):

$$\begin{cases} G \geq G^\mathcal{T} \\ t \leq N \end{cases}$$

are simplified, and the task is considered successfully finished if the units of work output by the team reach the minimum  $G^\mathcal{T}$ :

$$G \geq G^\mathcal{T}$$

agent  $i$  is given a fixed weight  $\beta_i$  at the beginning of the game, so their contribution  $g_i$  is aggregated with the rest as:

$$G = \left( \sum_{i \in I} \beta_i \cdot g_i^\rho \right)^\frac{1}{\rho}$$

After this only turn, agent  $i$  receives a reward  $R_i$  that depends on the number of leisure units that  $i$  has been able to enjoy,  $x_i$ , and the assessment  $\sigma(G)$ , which is common to all agents. This reward is given by the utility function of the underlying teamwork game:

$$R_i = x_i^\alpha \cdot \sigma(G) \in \mathbb{R}$$

## 6.3. Learning agents

Our set of agents,  $I = \{1, \dots, n\}$ , consists of  $n$  independent  $k$ -armed bandits. Each agent  $i$  has an intrinsic, fixed expertise level  $p_i^\mathcal{T}$ . Following the convention established in Section 4, we assume all agents have a leisure capacity  $p_i^L = 1$  (although, in general, this parameter could be different from 1). Given a task  $\mathcal{T}$ , agent  $i$  must decide what percentage  $a_i \in \mathcal{A}_i$  of time  $\Delta_t$  to allocate to  $\mathcal{T}$ . The action vector  $\mathbf{a} = (a_1, \dots, a_n)$  represents the strategy profile of all players, which maps to a vector of contributions  $\mathbf{g} = (g_1, \dots, g_n)$ . Consequently, their private allocation to leisure will be

$$x_i = \Delta_t - \frac{g_i}{p_i^\mathcal{T}} = \Delta_t \cdot (1 - a_i),$$

where we have used the relationship  $g_i = a_i \cdot p_i^\mathcal{T} \cdot \Delta_t$  from Eq. (16) to express  $x_i$  in terms of  $a_i$ .

Notice that this equation is correct as long as the convention  $p_i^L = 1 \forall i$  is followed. We manage  $n$  independent  $k$ -armed bandits, each selecting actions from a Boltzmann distribution with a temperature parameter  $\tau > 0$ :

$$\mathbb{P}(a) = \frac{e^{Q_t(a)/(Q_t^{\max} \cdot \tau)}}{\sum_a e^{Q_t(a)/(Q_t^{\max} \cdot \tau)}}$$

Each agent can distinguish actions (time percentages) with a granularity of 2 decimals. That is to say: agents can choose among the  $10^2 + 1$  actions in  $\mathcal{A}_i = \{0, 10^{-2}, \dots, 1 - 10^{-2}, 1\}$ . At episode  $t + 1$ , the value estimates of agent  $i$  are computed incrementally:

$$Q_i^{t+1} = Q_i^t + l_r^t \cdot (R_i^t - Q_i^t) \quad (40)$$

Here,  $Q_i^t$  represents agent  $i$ 's value estimate at episode  $t$ , while  $R_i^t$  denotes the reward that  $i$  receives in episode  $t$ . The learning rate parameter, denoted by  $l_r^t \in (0, 1]$ , is the same for all agents and decays with the inverse of time:

$$l_r^t = \left( \frac{\kappa}{\kappa + t} \right)^{-1} \quad (41)$$

As justified in Section 2, this algorithm implements smooth best responses: if  $\tau$  is sufficiently low and the agents interact for a reasonably long period, then convergence will only occur at a Nash distribution [16].

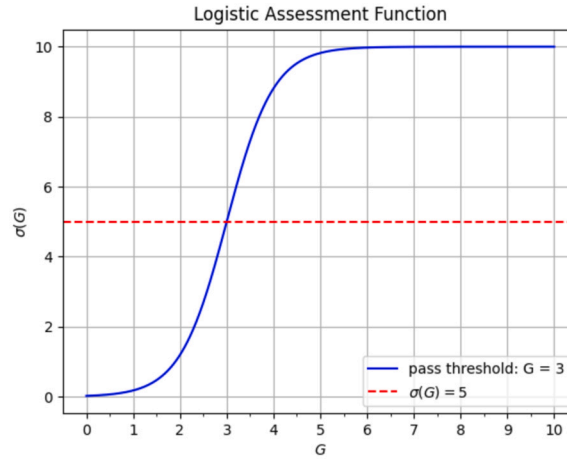


Fig. 4. Logistic evaluation function with passing threshold  $b = 3$ , steepness parameter  $\gamma = 2$ , and right-hand asymptote  $d = 10$ .

## 7. Results and discussion

The primary aim of this section is to address whether a MA-MAB framework, as described in Section 6, can converge to approximations of equilibria of teamwork games after learning. If convergence is possible, we will examine the nature of these empirical approximations and how they are impacted by changes in the various components of our model. We focus on teams consisting of two agents, leaving the study of larger teams for future work. More specifically, we conduct the following analyses:

1. We examine the consistency of the agents' learned strategies with the game-theoretic predictions of the underlying game's Nash Equilibria.
2. We examine the repercussions on agent team productivity stemming from the interplay between team composition and task types.
3. We explore how the difficulty of assessment influences agent team productivity.
4. We delve into the individual strategies adopted by agents and how they are impacted by different task types, team compositions, and assessment difficulties.
5. Finally, we expand our analysis beyond analytically solvable games and empirically study the policies of our agents after learning in a binary pass/fail scenario.

### 7.1. Experimental setup

Our experimental setup features teams composed of two multi-armed bandits, each equipped with 101 arms. These bandits utilise Boltzmann action selection with a low temperature parameter ( $\tau = 0.1$ ), similar to the value reported by [16]. Each arm represents a discrete value corresponding to the percentage of the decision-maker's total available time allocated to the task. We train each agent during  $5 \cdot 10^4$  episodes of 1 step each. Throughout the experiments, we consider that all of the agents' contributions are equally valuable ( $\beta_1 = \beta_2 = 1$ ) and assume a standard leisure time parameterization,  $p_1^L = p_2^L = 1$ . The learning rate parameter, denoted by  $l_r' \in (0, 1]$ , is the same for all agents and decays with the inverse of time as per Eq. (41), where  $k > 0$  is determined using an automatic hyperparameter optimization framework [82]. By default, we conduct one-shot games with a turn duration of  $\Delta_t = 10$  time units, and we set the common parameter  $\alpha$  in the utility function to 2 (Eq. (27)). This value of  $\alpha$  models agents for whom their free time is squaredly more important than the teamwork outcome. Both  $\Delta_t$  and  $\alpha$  are convenient choices we have made to ensure clarity in presenting the results, but in general, any  $\Delta_t, \alpha > 0$  could have been chosen. The reward function guiding agent  $i$ 's policy is derived from the underlying utility in the respective teamwork game:

$$R_i \equiv \tilde{u}_i(g_i, \sigma(G)) = x_i^\alpha \cdot \sigma(G) = \Delta_t^\alpha \cdot (1 - a_i)^\alpha \cdot \sigma(G)$$

Notice that  $\sigma(G)$  is the same for both agents, and the second equality follows from Eq. (19). As for the evaluation, we seek a function that can represent continuous assessment. Ideally, it should be readily applicable to provide a continuous representation of evaluations based on varying levels of merit. As discussed below, the logistic function (Fig. 4) meets these requirements:

$$\sigma(G) = \frac{d}{1 + e^{-\gamma(G-b)}} \quad (42)$$

The logistic function takes the teamwork outcome  $G$  as input and returns an evaluation within the interval  $[0, d] \subset \mathbb{R}$ . The parameter  $\gamma$  indicates how many units of the teamwork outcome are necessary to raise the assessment by one point. When  $G = b$ , the teamwork outcome receives an evaluation of  $\frac{d}{2}$ , making  $b$  the passing threshold. Notably, the logistic function allows us to create a continuous approximation of qualitative evaluations, similar to a gradable pass/fail system. Furthermore, by stacking logistic

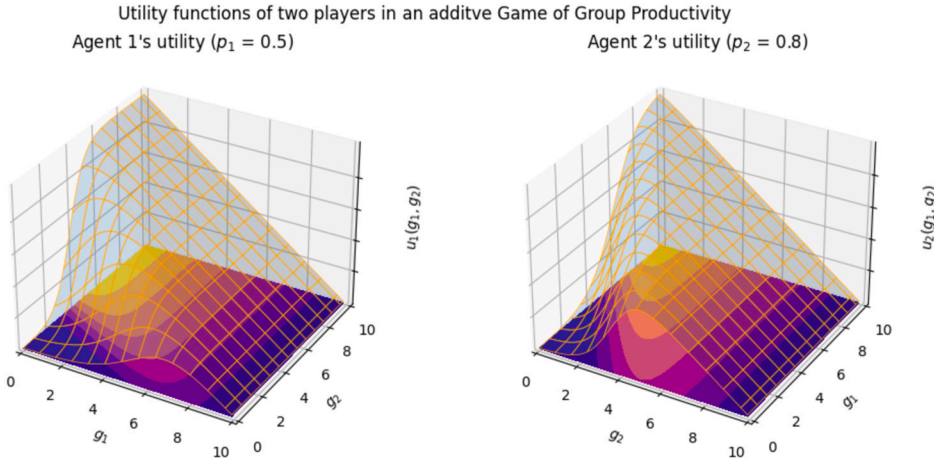


Fig. 5. Utilities  $\tilde{u}_i$  of two agents in a team, shown as functions of the agent's work units contribution  $g_i$  and the rest of the team's contribution  $G_{-i}$ . Left: Utility for an agent with expertise  $p_1^T = 0.5$ . Right: Utility for their teammate with expertise  $p_2^T = 0.8$ .

functions, we could achieve a continuous representation of evaluations based on varying levels of merit. Some of the key properties of the logistic function include:

- **Non-negative assessment:** The function approaches the right horizontal asymptote,  $d$ , representing the maximum attainable reward. The left asymptote is  $\sigma(G) = 0$ , although, in our model,  $G$  remains non-negative, preventing it from being reached.
- **Evaluation steepness:** The curve steepness, controlled by parameter  $\gamma$ , determines the rate of evaluation change concerning  $G$ .
- **Passing threshold:** The parameter  $b$  of the curve represents the midpoint of the evaluation criteria, and it determines the passing threshold.

As discussed in Section 5.3.1, the ratio  $\frac{\sigma(G)}{\sigma'(G)}$  indicates how valuable it is for players to invest their time in improving their assessment rather than enjoying their leisure time. In the case of the logistic function, the parameters  $b$  and  $\gamma$  jointly provide this information. If  $b$  decreases, then the passing threshold lowers, making it easier to pass (and making it less worthwhile to work on the task). If  $\gamma$  decreases, then the slope of the curve decreases and more teamwork outcome units are needed to increase the evaluation by one point (and also inhibiting contributions). Therefore,  $b$  and  $\gamma$  provide information about the *opportunity cost* of the evaluation in terms of enjoyed leisure. In this paper, we keep  $\gamma$  constant in all experiments for the sake of parsimony. Consequently, it will be the change in  $b$  from one experiment to another that marks differences in the fraction  $\frac{\sigma(G)}{\sigma'(G)}$ , and we will use the term *evaluation hardness* to refer to the effect of  $b$  in the following discussion.

The logistic function in Eq. (42) fulfils all the criteria to serve as an evaluation function (Definition 4.5). The resulting reward, expressed as a function of agent  $i$ 's contribution  $g_i$  and the provision of the teamwork outcome  $G$ ,  $\tilde{u}_i(g_i, \sigma(G))$ , remains non-negative across all values.

Fig. 5 depicts the utilities for each agent in a two-agent team: Agent 1 with  $p_1^T = 0.5$  and Agent 2 with  $p_2^T = 0.8$ . The logistic evaluation function is employed. The task is additive ( $\rho = 1$ ) with a passing threshold  $b = 5$  and turn duration  $\Delta_t = 10$ . It's worth noting that the most favourable scenario for Agent 1 occurs when their teammate contributes maximally ( $g_2 = 10$ ) while they refrain from any action ( $g_1 = 0$ ). When none of the agents contributes ( $g_1 = g_2 = 0$ ), both agents perceive a utility close to zero due to the evaluation function. However, neither agent desires to contribute maximally, as it would mean sacrificing leisure time and receiving a null utility.

For the specific case depicted in Fig. 5, the equilibrium aggregates of individual contributions (in work units) are ( $g_1^* = 1.45, g_2^* = 4.48$ ).

## 7.2. Consistency with game-theoretic predictions

First, we scrutinise the alignment of our agents' learned strategies with the NE of the underlying teamwork game. A total of 140 experiments are conducted, involving pairs of agents with diverse expertise combinations drawn from a set ranging from  $p_i^T = 0.1$  to  $p_i^T = 0.9$ . Additionally, we explore different permutations of task types, ranging from strongly conjunctive tasks ( $\rho = -100$ ) to strongly disjunctive tasks ( $\rho = 100$ ) (Table 5).

We use three evaluation conditions represented by thresholds  $b \in \{3, 5, 7\}$  referred to as **soft**, **medium**, and **hard** assessment treatments, respectively (Table 6).

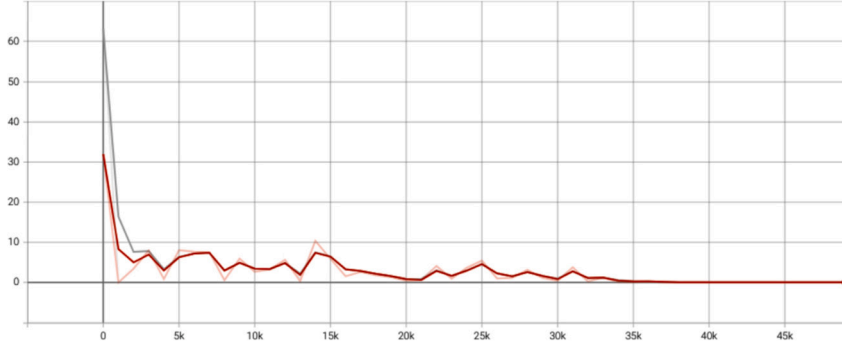
For each experimental configuration  $[(p_1, p_2), \rho, b]$ , we train the pair of agents. Each training consists of playing a teamwork game  $5 \times 10^4$  times. During these episodes, agents learned a Q-value function following the algorithm in Eq. (40). The convergence of the

**Table 5**Task types used in the experiments, characterised by the value of  $\rho$ .

Strongly Conjunctive	Quasi-Conjunctive	Additive	Quasi-Disjunctive	Strongly Disjunctive
$\rho = -100$	$\rho \in \{-10, -3, 0.5\}$	$\rho = 1$	$\rho \in \{3, 10\}$	$\rho = 100$

**Table 6**Evaluation conditions used in the experiments, represented by passing thresholds  $b \in \{3, 5, 7\}$ .

Soft	Medium	Hard
$b = 3$	$b = 5$	$b = 7$

**Fig. 6.** Representation of the convergence of  $Q$ -values during training of two agents. Timesteps are represented in the horizontal axis and the difference  $\max_a \{Q_{n+1}(a) - Q_n(a)\}$  is represented in the vertical axis.

agents was quantitatively assessed by monitoring the stability of the  $Q$ -table. When the changes in the  $Q$ -values stabilised below a predefined threshold  $\delta$  across episodes:

$$\text{Convergence Criterion: } \max_a |Q^{t+1}(a) - Q^t(a)| \leq \delta \quad (43)$$

we concluded that the learning process had reached convergence. To evaluate the learning dynamics of the agents, we visualised the  $\text{argmax}$  of the  $Q$ -function over time, which represents the actions that the agents believe to be optimal. Fig. 6 illustrates the evolution of these optimal actions for two agents throughout the training process.

The learned policies for each experimental configuration,  $\tilde{\mathbf{a}} = (a_1, a_2)$ , result in an experimental teamwork outcome  $\tilde{G}$ . This experimental result is compared with the theoretical aggregate of the teamwork outcome at the Nash equilibrium,  $\hat{G}$ , computed according to Sections 5.3.1, 5.3.2, and 5.3.3 respectively. Fig. 7 shows the scatter plot and the density estimation contour for these pairs  $(\tilde{G}, \hat{G})$ .

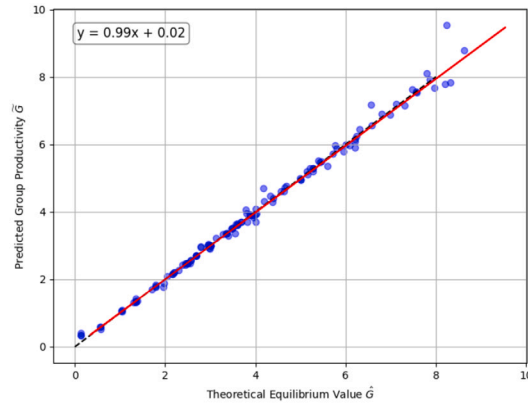
Overall, these results underscore the robustness of our multiagent system, as evidenced by the remarkable agreement between our agents' learned strategies and the corresponding NE strategies. This alignment validates the efficacy of our learning framework to approximate the Nash equilibria of teamwork games. Furthermore, the near-perfect fit observed between the theoretical and empirical values suggests that our system is well-designed and theoretically grounded, paving the way for its application in diverse real-world scenarios beyond the scope of traditional game theory assumptions. On the left side of Fig. 7, we present a scatter plot depicting the relationship between the theoretical equilibrium aggregate  $\hat{G}$  and the empirical team productivity  $\tilde{G}$  derived from the MA-MAB setting. The horizontal axis represents the theoretical equilibrium aggregate  $\hat{G}$ , while the vertical axis represents the empirical or predicted team productivity. Additionally, the plot includes the identity line  $y = x$  (dashed black line) and a regression line (shown in red). The parameters of the linear regression are  $(\alpha = 0.99, \beta = 0.02)$ , with a goodness of fit of  $\xi^2 = 0.992$ .

Since the temperature parameter governing the smooth best responses is maintained at a low level ( $\tau = 0.1$ ), the agents' strategies approximate Nash equilibria [16]. Consequently, the aggregate values that we represent can be considered equilibrium aggregates, directly derived from the converged strategies of the agents. This convergence to values near the equilibrium aggregate implies that our agents' learned policies are also near-equilibrium strategies. On the other hand, the density estimation contours displayed on the right side of the figure highlight the regions where the majority of the probability mass is concentrated. These contours provide evidence of the consistency between our agents' learned strategies and the NE predictions.

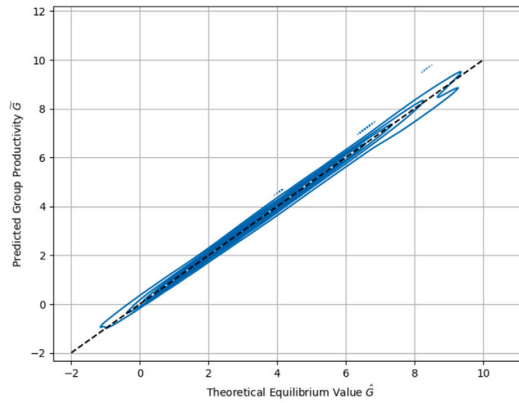
### 7.3. Interaction between team composition and task type

In this section, we aim to highlight our main observations regarding how task type affects team productivity. We will use Fig. 8 to illustrate our findings.





(a) Scatter plot



(b) Density estimation contour

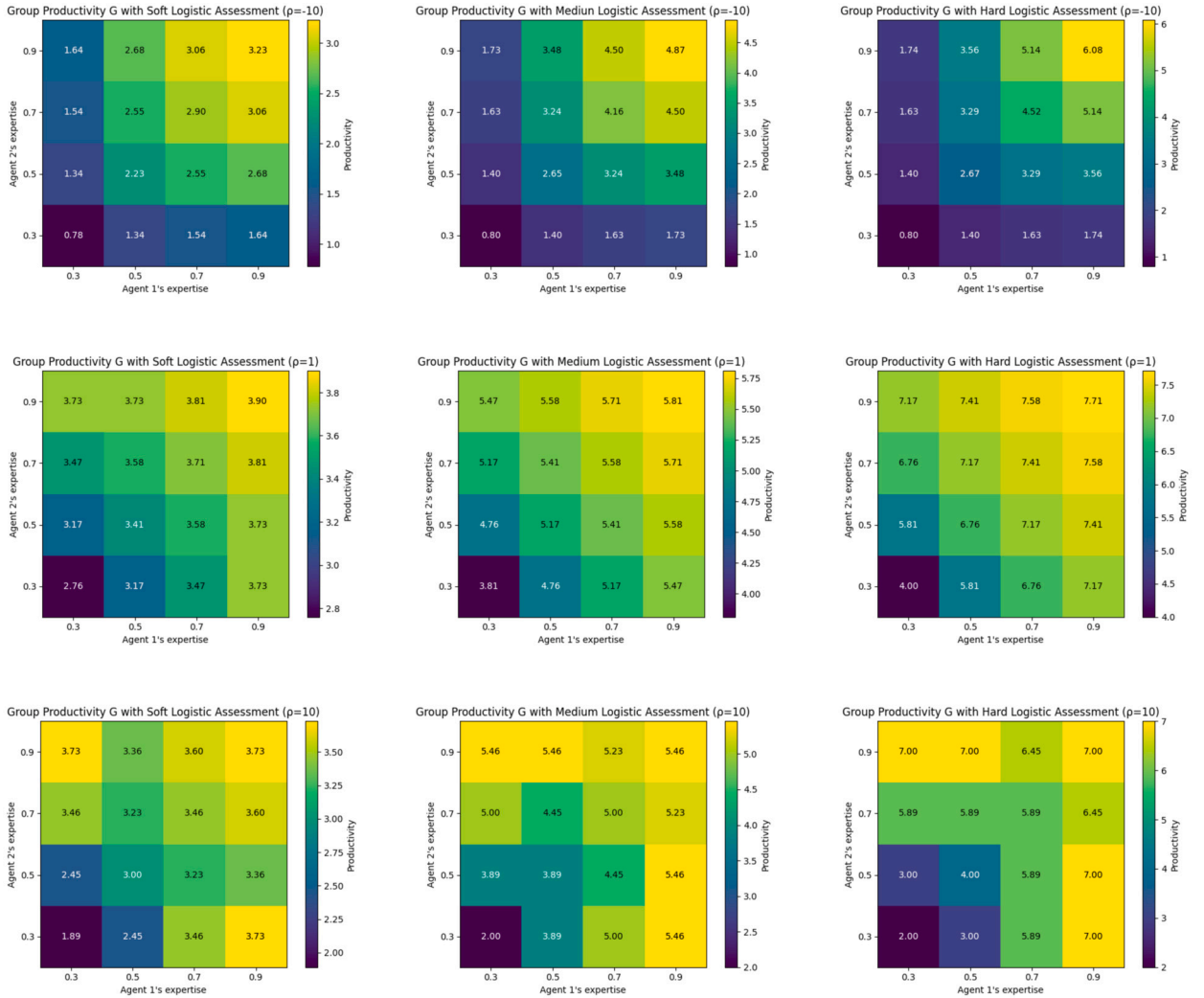
**Fig. 7.** (a) Scatter plot showing the correspondence between theoretical Nash equilibrium (NE) aggregate productivity values (horizontal axis) and the empirical team productivity (vertical axis) derived from the MA-MAB setting. The dashed black line represents the identity line  $y = x$ , while the red line shows a linear regression fit with parameters  $\alpha = 0.99, \beta = 0.02$  and a goodness of fit of  $\xi^2 = 0.992$ . This close alignment underscores that, upon convergence, the aggregate values closely approximate those expected under NE, suggesting that the agents' learned strategies are near-equilibrium policies. (b) Density estimation contour highlighting regions of concentrated probability mass, which further demonstrates the consistency between the agents' learned strategies and NE predictions.

This figure utilises a colour scale to represent the levels of team productivity predicted by our model for different teams. We consistently compare the same teams, formed by all possible combinations of pairs of agents with expertise levels extracted from  $p_i^T \in \{0.3, 0.5, 0.7, 0.9\}$ , and create a plot for quasi-conjunctive, additive, and disjunctive tasks ( $\rho \in \{-10, 1, 10\}$ ). We consider passing thresholds from soft to high:  $b \in \{3, 5, 7\}$ . The current discussion emphasises the changes in team productivity as the task type varies. Thus, we will be describing how team productivity changes as one moves **across rows** in the figure.

Focusing on conjunctive tasks (first row of Fig. 8), we observe that as one moves perpendicularly away from the diagonal of homogeneous teams, performance drops. For example, there is a decrease of 38% (from  $G = 2.65$  to  $G = 1.63$ ) in team productivity between the homogeneous pair  $\mathbf{p} = (0.5, 0.5)$  and the heterogeneous pair  $\mathbf{p} = (0.3, 0.7)$  in the conjunctive task with medium assessment treatment. We interpret this fact as a confirmation of Steiner's assertion that **conjunctive tasks are potentially better carried out by homogeneous teams** (with similar expertise levels) than heterogeneous ones [35], at least in the case of rational agents operating under the utility functions considered here.

Regarding additive tasks (middle row), we observe that they consistently yield the highest outcomes no matter what column (assessment treatment) we look at. This aligns with findings in social psychology, where additive tasks typically result in higher team potential than the best team member's potential. Visually, team productivity remains constant or nearly constant for a given passing threshold along the direction perpendicular to the identity diagonal of the represented matrices. For example, pairs with expertise levels ( $p_i^T = 0.5, p_j^T = 0.7$ ) and ( $p_i^T = 0.3, p_j^T = 0.9$ ), exhibit very similar performances among themselves. Again, we consider this a confirmation of Steiner's hypothesis that **team productivity is unaffected by the heterogeneity of team members in additive tasks** [35].

Finally, the productivity observed in the disjunctive task falls somewhere between that observed in the conjunctive and additive tasks across columns. A visual examination of the bottom row of Fig. 8 reveals that, unlike the conjunctive scenario, team performance



**Fig. 8.** Variation in team productivity across different teams for 1) conjunctive tasks (upper row,  $p = -10$ ), 2) additive tasks (middle row,  $p = 1$ ), and 3) disjunctive tasks (lower row,  $p = 10$ ), dependent on the passing threshold  $b$  (left column:  $b = 3$ , middle column:  $b = 5$ , right column:  $b = 7$ ). Colour legend: red indicates low productivity, while green indicates high productivity, with the passing threshold determining the colour transition.

$G$  declines as one moves perpendicularly to the diagonal of homogeneous teams. For instance, consider the pairs with expertise levels  $\mathbf{p} = (0.5, 0.5)$  and  $\mathbf{p} = (0.3, 0.7)$  in the disjunctive task with medium assessment treatment. The first team produces  $G = 3.89$  units of teamwork outcome, whereas the second team produces  $G = 5.00$ . This results in an increase in productivity of 29%. This observation reinforces Steiner's proposition that **disjunctive tasks may be more effectively tackled by teams with diverse expertise levels** rather than by those with uniform capabilities [35].

#### 7.4. Interaction between team composition and assessment treatment

Based on our theoretical analyses in Section 5.3, the difficulty of the assessment should influence productivity. In this section, we investigate whether this impact is observed in the empirical approximations of equilibrium teamwork outcomes achieved by our agents. Given a task, we can study how productivity varies from a soft assessment treatment to a hard assessment treatment as we adjust the parameter  $b$  in the logistic function (Eq. (42)). To make these comparisons, we look again at Fig. 8 and turn our attention to changes in  $G$  **column-wise**. Take for instance the second row (additive task) in Fig. 8. In the soft assessment condition (left column), all teams except one, comprising two agents with low expertise ( $p = 0.3$ ), pass. However, as the pass threshold is raised, more teams fail. In the hard assessment condition, only teams with an aggregated expertise of at least 60% of the total possible expertise pass:

$$\frac{\sum_{j=1}^n p_j^T}{n} \geq 60\%$$

Given a fixed team, productivity generally increases as the pass threshold is raised. However, these increments are consistently smaller when resulting productivity falls short of the new passing threshold. A consistent behavioural pattern emerges across different task types: when the task is easy enough, productivity increases as  $b$  is raised, and it increases even more when passing becomes possible at the next level of evaluative difficulty. This trend ceases once the task requirements exceed the team's capacity

While it's evident that teams typically exhibit poorer performance on conjunctive tasks (top row), this does not imply an absence of potential improvement. Indeed, our plots reveal that when agents have the opportunity to enhance their performance (*i.e.*, when their expertise levels allow for increased contributions without significant sacrifices in their free time), more rigorous evaluation criteria result in higher team productivity values during conjunctive tasks. However, when agents have sufficiently low expertise levels, this incentive mechanism becomes futile, as the cost of increasing contributions outweighs the potential benefits in terms of their free time. The scenario depicted on the upper row of Fig. 8 illustrates an intrinsically challenging task, where under the soft assessment treatment, only the top two teams ( $\mathbf{p} = (0.7, 0.9)$  and  $\mathbf{p} = (0.9, 0.9)$ ) manage to succeed. To enhance performance, it may be necessary to provide agents with additional time  $\Delta_i$ , or to soften the assessment by reducing the evaluation threshold  $b$ .

As for disjunctive tasks (bottom row), we once again notice a trend towards increased team productivity as the evaluation criteria become more stringent. The nature of the disjunctive task alters the threshold that defines what constitutes “low” team expertise compared to the conjunctive scenario. In a conjunctive task, hardening the assessment with an increase of  $b$  from 3 to 5 barely elicited a response in productivity levels from pairs of agents whose maximum expertise was less than or equal to 0.5. For example, the pair  $\mathbf{p} = (0.3, 0.3)$  only raised their teamwork outcome from 0.78 to 0.8. With a further increase of  $b$  from 5 to 7, the agents became even more insensitive to the evaluation threshold, with only teams boasting minimum expertise of  $p_i^T \geq 0.7$  showing an increase in contributions by at least one-tenth. For example, the pairs with  $\mathbf{p} = (0.7, 0.9)$  raised their teamwork outcome from 4.5 to 5.14, a 14%. However, the pairs  $\mathbf{p} = (0.5, 0.7)$  only increased their outcome by 1%. In the disjunctive setting, all teams react to the initial increase in assessment hardness, and those with at least one member with high expertise  $p_i^T \geq 0.7$  respond to the subsequent difficulty escalation. Notably, we observe a decrease in team productivity during the transition from “medium” to “hard” for the pairs with low expertise levels ( $p_{1,2}^T \leq 0.5$ ). Thus, in this case, an increase in the demand threshold has a detrimental effect on team productivity, a **phenomenon akin to a loss of motivation in human teams [35] that can be explained as a direct consequence of our agents' reward-maximising behaviour.**

Finally, we observe that as the task becomes more challenging, the amount by which productivity improves, compared to the previous difficulty level, becomes smaller. In other words, while productivity still tends to increase with increasing difficulty, the rate of this increase diminishes. Specifically, for the studied additive, conjunctive and disjunctive tasks, we consistently observe a reduction in the rate at which the output productivity  $G$  increases as the assessment becomes more stringent (*i.e.*, with higher  $b$ ).

### 7.5. Individual strategies

The previous sections have focused on the team outcome  $G$  and how it changes as the assessment criteria become more demanding or the nature of the task is altered. Now we study how the assessment difficulty and the task type affect the **individual** strategies followed by the agents. As a reminder, by *strategies* we refer to the percentage  $a_i$  of their turn that the agents choose to allocate to the task. An agent with expertise  $p_i^T$  who takes strategy  $a_i$  makes a teamwork contribution  $g_i$  (Eq. (16)). Table 7 presents the learned strategies teams of two agents for an additive task ( $\rho = 1$ ), where  $\Delta_i = 10$  and  $\alpha = 2$  for all agents. The table showcases strategies for the pairs who are assigned an additive task, categorised under three assessment treatments: soft (lighter-shaded), medium (medium-shaded), and hard (darker-shaded). For simplicity, we fill only the upper diagonal of the matrices in the table, which are symmetric.

We consistently observe in this table that the agent with the highest expertise level contributes more time than their counterpart in all pairs. Remarkably, elevating the evaluation level does not diminish individual contributions in any observed case, albeit there are instances of no discernible effect. For instance, the less proficient agent in the ( $p_1^T = 0.3, p_2^T = 0.9$ ) pair contributes nothing under soft or medium evaluation (free-riding), only engaging with a 20% commitment when the passing threshold reaches  $b = 7$ .

Still in the additive case, the average contribution of an agent with expertise  $p_i^T$  increases across all teams as the passing threshold rises. However, transitioning from a soft to a medium evaluation system yields a more pronounced effect than transitioning from medium to hard. In other words, agents consistently increase their task engagement as evaluation demands intensify, with a more significant augmentation observed when transitioning from a soft to medium evaluation compared to a medium to hard evaluation.

For further understanding of individual strategies in equilibrium for conjunctive and disjunctive tasks, readers can refer to similar Tables F.10, F.11 like Table 7 in Appendix F. In the conjunctive scenario, we notice that the weakest member of all possible pairs in all assessment treatments contributes more effort than their counterpart. This phenomenon is akin to the **Köhler effect** observed in human teams, where an inferior team member performs a difficult task better in a team or cooperative situation than expected based on their individual performance [83]. Furthermore, increasing the evaluation level did not decrease individual contributions in any observed case, although there were some instances where there was no discernible effect.

On the other hand, in the disjunctive case, elevating the evaluation level causes a drop in individual contributions when the expertise of the strongest member is sufficiently high ( $> 0.7$ ). For instance, if we look at the corresponding table in Appendix F, we can see that the less proficient agent in the ( $p_1 = 0.5, p_2 = 0.7$ ) pair contributes in one of the equilibria under soft evaluation, but free-rides as soon as the passing threshold is heightened. This phenomenon mirrors the **social loafing** effect observed in human teams, where some members exploit the efforts of others while free-riding.

Table 8 displays the percentage increase in the average team contribution of each agent to the additive, conjunctive and disjunctive task as the passing threshold was raised from soft to medium and from medium to hard.

**Table 7**

Individual team contribution across different teams for an additive task under three different assessment treatments. For simplicity, we fill only the upper diagonal of the matrices, which are symmetric.

(a) Soft Assessment Treatment ( $b = 3$ ).				
$p_1^T \backslash p_2^T$	0.30	0.50	0.70	0.90
0.30	(46%, 46%)	(20%, 52%)	(0%, 50%)	(0%, 41%)
0.50		(34%, 34%)	(16%, 40%)	(0%, 41%)
0.70			(27%, 27%)	(23%, 22%)
0.90				(22%, 22%)

(b) Medium Assessment Treatment ( $b = 5$ ).				
$p_1^T \backslash p_2^T$	0.30	0.50	0.70	0.90
0.30	(64%, 64%)	(46%, 68%)	(20%, 66%)	(0%, 61%)
0.50		(52%, 52%)	(34%, 53%)	(16%, 53%)
0.70			(40%, 40%)	(27%, 43%)
0.90				(32%, 32%)

(c) Hard Assessment Treatment ( $b = 7$ ).				
$p_1^T \backslash p_2^T$	0.30	0.50	0.70	0.90
0.30	(67%, 67%)	(64%, 78%)	(46%, 77%)	(20%, 73%)
0.50		(68%, 68%)	(52%, 66%)	(34%, 63%)
0.70			(53%, 53%)	(40%, 53%)
0.90				(43%, 43%)

**Table 8**

Percentage increase in average dedication ( $\Delta_{\%}(\bar{a}_i)$ ) of each agent to different tasks as the passing threshold rises.

(a) Additive Task ( $\rho = 1$ ).		
$p_i^T$	Soft ( $b = 3$ ) to Medium ( $b = 5$ ) $\Delta_{\%}(\bar{a}_i)$	Medium ( $b = 5$ ) to Hard ( $b = 7$ ) $\Delta_{\%}(\bar{a}_i)$
0.3	97%	52%
0.5	104%	50%
0.7	68%	39%
0.9	45%	35%

(b) Conjunctive Task ( $\rho = -10$ ).		
$p_i^T$	Soft ( $b = 3$ ) to Medium ( $b = 5$ ) $\Delta_{\%}(\bar{a}_i)$	Medium ( $b = 5$ ) to Hard ( $b = 7$ ) $\Delta_{\%}(\bar{a}_i)$
0.3	5%	0%
0.5	27%	1%
0.7	46%	11%
0.9	53%	24%

(c) Disjunctive Task ( $\rho = 10$ ).		
$p_i^T$	Soft ( $b = 3$ ) to Medium ( $b = 5$ ) $\Delta_{\%}(\bar{a}_i)$	Medium ( $b = 5$ ) to Hard ( $b = 7$ ) $\Delta_{\%}(\bar{a}_i)$
0.3	6%	0%
0.5	4%	-36%
0.7	42%	18%
0.9	49%	28%

In all cases but one ( $\rho = 10$ ,  $p_i^T = 0.5$ ), the percentage increase as the evaluation hardens is non-negative. That is, the more demanding the evaluation, the more committed the agents. In the mentioned exception, the weakest agent learns not to contribute as soon as the task becomes maximally difficult (hard treatment) and their counterpart possesses greater expertise, hence the observed negative difference in the table. These increments tend to be higher overall when agents face an additive task compared to other cases, although this is not always the case (the agent with expertise  $p_i^T = 0.9$  always learns to lead regardless of the task type, exhibiting a steady pattern and consistently increasing its contributions by a similar amount for a given assessment treatment).

Observing Table 8, we discern that agents with lower expertise levels demonstrate a heightened sensitivity to the type of task at hand. For example, in the case of  $p_i^T = 0.3$  in an additive task, the agent's effort nearly doubles when the assessment is hardened and the passing threshold shifts from  $b = 3$  to  $b = 5$ . Examining Table 7a, this notable escalation stems from the agent's non-contribution under the soft assessment treatment ( $b = 3$ ) when their counterpart exhibits higher expertise ( $p_2^T = 0.7$ ). Likewise, transitioning from medium to hard assessment treatment prompts this agent to contribute even when their counterpart possesses substantial expertise

( $p_2^T = 0.9$ ). In conjunctive tasks ( $\rho = -10$ ), this agent consistently contributes, leaving minimal scope for performance improvement. Finally, in disjunctive tasks ( $\rho = 10$ ), the agent refrains from contributing altogether in any equilibrium where their counterparts surpass them significantly in expertise (i.e.,  $p_2^T > 0.5$ ). This equilibrium remains resilient to variations in assessment treatment, indicating that **hardening the assessment does not induce changes in this agent's policy**. In general, agents with high expertise ( $p_i^T \geq 0.7$ ) exhibit more steadfast policies concerning task types. Regardless of the task, they consistently enhance their contributions as the passing threshold rises, with a comparable degree of increase across all tasks. Hence, these agents are more influenced by shifts in evaluation criteria than by the inherent nature of the task.

### 7.6. A binary pass/fail assessment scenario

To explore the versatility of our MA-MAB simulation platform beyond game-theoretical contexts, we explore a scenario featuring a discontinuous reward function, mirroring a binary pass/fail evaluation scenario. Our reward function is now defined as:

$$R_i = \Delta_i^\alpha \cdot (1 - a_i)^\alpha \cdot H(G) \quad (44)$$

where  $H(G)$  is a modified version of the Heaviside function (Fig. 9):

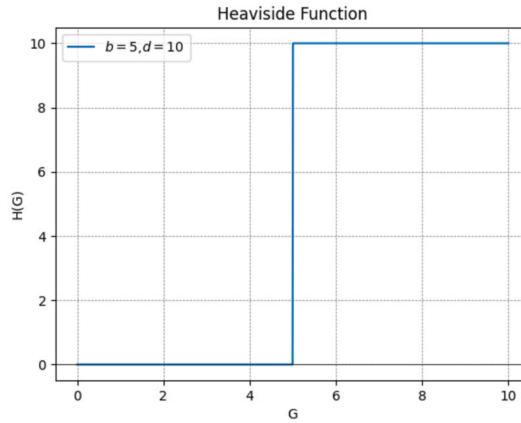


Fig. 9. A Heaviside function with parameters  $b = 3$  and  $d = 10$ .

The function returns 0 for  $G < b$  and  $d$  for  $G \geq b$ . In other words, it acts as a step function that “switches on” at  $G = b$ :

$$H(G) = \begin{cases} 0 & \text{if } G < b \\ d & \text{if } G \geq b \end{cases}$$

Note how  $d$  here assumes the role of the right horizontal asymptote in Eq. (42), while  $b$  represents the passing evaluation point. To maintain consistency with previous sections, we set the maximum evaluation possible  $d = 10$ . We consistently compare the same teams as in Section 7.3, formed by all possible combinations of pairs of agents with expertise levels extracted from  $p_i^T \in \{0.3, 0.5, 0.7, 0.9\}$ . We concentrate on additive tasks under the medium assessment treatment ( $\rho = 1$ ,  $b = 5$ ). For each experimental configuration  $[(p_1^T, p_2^T), \rho = 1, b = 5]$ , we train the pair of agents, with each training consisting of playing a teamwork game with a Heaviside evaluation function  $5 \times 10^4$  times. The learned policies for each experimental configuration,  $\tilde{\mathbf{a}} = (a_1, a_2)$ , result in an experimental teamwork outcome  $\tilde{G}$ . Similar to Section 7.3, we employ a colour scale to depict teamwork outcome under this discontinuous reward regime (Fig. 10). Red signifies lower productivity levels, while green denotes higher productivity levels. To compute the displayed quantities in the figure, each game was learned three times, resulting in three values of the teamwork outcome for each team  $\mathbf{p}$ ,  $\{\tilde{G}_1^{\mathbf{p}}, \tilde{G}_2^{\mathbf{p}}, \tilde{G}_3^{\mathbf{p}}\}$ . From each sample, we calculate the mean  $\bar{G}$ . Fig. 10 shows the resulting  $\bar{G}$  for all teams under this binary pass/fail assessment scenario. To ensure the accuracy of average value  $\bar{G}$ , we verified that the percentage dispersion of  $\tilde{G}$  in each sample was less than 2%, where the percentage dispersion is computed as:

$$\text{Percentage Dispersion of } G = \frac{\max\{\tilde{G}_1^{\mathbf{p}}, \tilde{G}_2^{\mathbf{p}}, \tilde{G}_3^{\mathbf{p}}\} - \min\{\tilde{G}_1^{\mathbf{p}}, \tilde{G}_2^{\mathbf{p}}, \tilde{G}_3^{\mathbf{p}}\}}{\bar{G}} \times 100\%$$

Table 9 presents the percentage dispersion of the samples  $\{\tilde{G}_1^{\mathbf{p}}, \tilde{G}_2^{\mathbf{p}}, \tilde{G}_3^{\mathbf{p}}\}$  for each team. For simplicity, we fill only the upper diagonal of the table, which is symmetric.

We observe that the strategies agents converge to when working on an additive task under the medium assessment treatment show some variability. However, the final team productivity  $G$  consistently remains the same across all experiments for the pair, with very low dispersion. A notable detail we have observed is that when pairs are highly heterogeneous (i.e. agents show large differences in expertise), particularly in cases such as  $(p_1^T = 0.3, p_2^T = 0.7)$ ,  $(p_1^T = 0.3, p_2^T = 0.9)$ ,  $(p_1^T = 0.5, p_2^T = 0.9)$ , the less skilled agent refrains from contributing. Interestingly, in this case, their behaviour is consistent, with variance again falling below 2%. This contrasts with

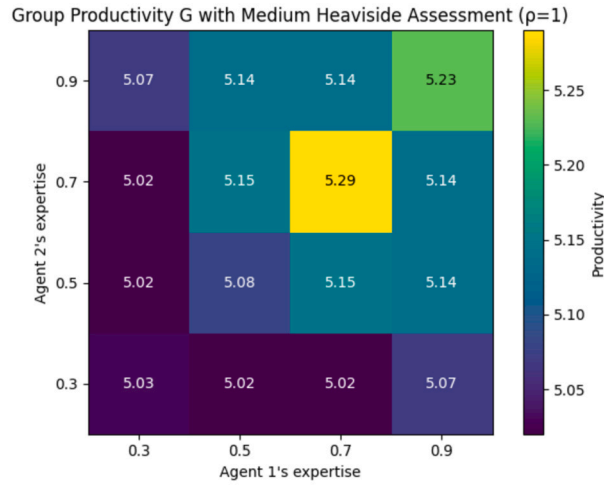


Fig. 10. Team productivity for an additive task under the medium assessment ( $b = 5$ ) treatment and using a Heaviside evaluation function. Colour Legend: Red would denote low productivity under the passing threshold, while Green signifies high productivity over the passing threshold.

Table 9

Percentage dispersion of the teamwork outcome samples  $\{\tilde{G}_1^p, \tilde{G}_2^p, \tilde{G}_3^p\}$  for each team. In all cases, the percentage dispersion is lower than 2%.

$p_i^T \backslash p_j^T$	0.30	0.50	0.70	0.90
0.30	0%	0.4%	0.6%	1.19%
0.50		1%	1.78%	0.2%
0.70			0.2%	0.18%
0.90				0%

the continuous assessment scenario depicted in Table 7b, where such behaviour was observed only in the cases of the two maximally heterogeneous pairs ( $p_i^T = 0.3, p_j^T = 0.9$ ). Therefore, the inclusion of a binary pass/fail evaluation function causes agents to polarise in their behaviour compared to continuous evaluation scenarios. In situations where there are significant differences between them (such as when the expertise gap is 0.4 or greater), the less skilled agent chooses to withdraw, thereby leaving their counterpart to shoulder the entire workload. Nonetheless, these adapted strategies ensure the team consistently meets the assignment criteria, as indicated by the absence of red hues in Fig. 10.

## 8. Conclusions, limitations, and future work

Game theory and computer science have extensively studied the concept of teamwork. In both fields, a team is traditionally perceived as a set of players or agents that need to cooperate with each other. Game theory assumes that all players in a team share the same utility function and examines teamwork using cooperative game theory tools. On the other hand, computer science, particularly Multiagent Systems (MAS), views teams as groups of self-interested agents who require cooperation to achieve individual success, rendering intentionally defective strategies unviable. Although both approaches have proven useful for studying environments where cooperation is in the interest of all players or agents, they are inadequate for modelling human teamwork.

Humans are capable of collaboration and recognise its benefits. However, collaboration may not be in the interest of all members of a human team, and establishing binding collaboration agreements is often unattainable in this context. Consequently, defection occurs.

This paper positions itself at the intersection of game theory and multiagent systems, aiming to address the question of how to learn theoretically grounded predictions of team performance in settings where cooperation is not enforced.

To answer this, we have dedicated the first part of this paper to formalising the concept of an aggregative teamwork game: a non-cooperative game that incorporates the elements identified by Steiner in social psychology as most informative for explaining the productivity of a human team. After presenting this model, we characterised the Nash equilibria (NE) of teamwork games. Notably, these equilibria extend beyond those seen in general public good games, with variations arising from the novel elements introduced in our model.

The second part of the article proposes a multiagent multi-armed bandit (MA-MAB) framework in which agents learn strategic behaviour that approximates the NE of the game. We validated the proposed MA-MAB system by empirically proving the convergence of the learned strategies towards the approximated NE of the game. After this validation, we analysed the impact of isolated variables on the teamwork outcome.



When we studied team productivity across different task types, our findings aligned with [Steiner](#)'s hypotheses on team composition and productivity. Specifically, conjunctive tasks where the weakest members have the most influence are better performed by homogeneous teams (similar expertise levels). In additive tasks (where contributions are summed), team composition does not affect productivity. Conversely, when the strongest members have the most influence (disjunctive tasks), heterogeneous teams (varying expertise levels) are more effective than homogeneous ones.

In analysing the effects of increasing the difficulty of the evaluation on teamwork outcome, we find that if the task is easy enough, team productivity increases. However, when the evaluation becomes so difficult that the task requirements exceed the capacity of the team, we observe that what [Steiner](#) would call a loss of motivation is actually an NE in the teamwork game, to which our reward-maximising agents' strategies converge.

When examining how individual strategies were impacted by task type and assessment difficulty, we found that agents with lower expertise adapted their strategies based on task type, but their contributions varied less with assessment difficulty. In conjunctive tasks, these agents contributed more compared to the other types of tasks, which mirrors the Köhler effect in human teams. Conversely, in additive and disjunctive tasks, the least skilled agent contributed less, sometimes completely free-riding. This phenomenon mirrors the social loafing effect in human teams. Meanwhile, agents with high expertise levels were more responsive to changes in evaluation criteria, contributing more as assessment difficulty increased. With equal assessment difficulty, high-expertise agents behaved nearly the same in all tasks.

In our last experiment, we expanded our analysis beyond analytically solvable games and empirically studied the policies of our agents after learning in a binary pass/fail scenario. Our findings showed that when pairs of individuals had high levels of heterogeneity, less skilled individuals tended to refrain from contributing, leaving their partner to handle the entire workload.

Our approach has some limitations. For example, we have only considered teamwork as a single interaction, while real-world teamwork typically involves multiple interactions over time. Furthermore, we have assumed that the parameter  $\alpha$ , which measures the importance of agents' free time, is the same for all agents. In the future, it would be interesting to explore how agents' policies change when  $\alpha$  is randomly distributed across the population. Similarly, the learning rate is currently the same for all agents and decays at the same pace, which may not be optimal [84]. Future studies will explore scenarios where parameter sharing is implemented judiciously rather than indiscriminately. We also plan to expand our simulations to include more agents, study the effects of team size on productivity, and incorporate discretionary tasks that have different weights for each agent's contribution.

In conclusion, our work has paved the way for a new research direction in cooperative AI [29], particularly in multiagent systems. This approach aims to achieve a better understanding of the social dynamics that drive collective behaviour in teams where cooperation is not enforced. We believe that incorporating factors that enhance the ecological validity of our model will enable us to provide a valuable tool for designing teamwork settings that favour cooperation in teams and societies in general.

## CRediT authorship contribution statement

**Alejandra López de Aberasturi Gómez:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Carles Sierra:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Jordi Sabater-Mir:** Writing – review & editing, Supervision, Methodology, Investigation, Funding acquisition.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Proof of Lemma 4.1

In a teamwork game, player  $i$ 's utility is determined by their private contribution  $x_i$  and the assessment that the team outcome  $G$  receives:

$$\tilde{u}_i(x_i, \sigma(G)) \tag{A.1}$$

Let  $\tilde{u}_i(x_i, G)$  be the utility function of some general public good game. The properties outlined in Definition 4.5 ensure that the new utility function  $\tilde{u}_i(x_i, \sigma(G))$  is both differentiable and quasi-concave:

- **Differentiability:** Let  $\sigma(G)$  and  $\tilde{u}_i(x, G)$  be both differentiable. For  $f(x, G) = \tilde{u}_i(x_i, \sigma(G))$  to be differentiable, we need to ensure that its partial derivatives are well-defined.
  - Since  $\sigma$  does not depend on  $x$ , the partial derivative of  $\tilde{u}_i(x, \sigma(G))$  with respect to  $x$  is the partial derivative of  $\tilde{u}_i(x, G)$  evaluated at  $(x, \sigma(G))$ :

$$\frac{\partial \tilde{u}_i(x, \sigma(G))}{\partial x} = \frac{\partial \tilde{u}_i(x, G)}{\partial x} \Big|_{(x, \sigma(G))}.$$

– Applying the chain rule and considering that  $\sigma(\cdot)$  is differentiable, we have:

$$\frac{\partial \tilde{u}_i(x, \sigma(G))}{\partial G} = \frac{\partial \tilde{u}_i(x, G)}{\partial G} \Big|_{(x, \sigma(G))} \cdot \sigma'(G).$$

Since both partial derivatives exist and are continuous,  $f(x, G) = \tilde{u}_i(x, \sigma(G))$  is differentiable.

- **Quasi-Concavity:** Let  $\tilde{u}_i(x, G)$  be a quasi-concave function and  $\sigma(G)$  be an increasing function. We want to show that  $f(x, G) = \tilde{u}_i(x, \sigma(G))$  is quasi-concave. Consider two arbitrary points  $(x_1, G_1)$  and  $(x_2, G_2)$  and a  $\lambda \in [0, 1]$ . The convex combination of these points is

$$(\lambda x_1 + (1 - \lambda)x_2, \lambda G_1 + (1 - \lambda)G_2).$$

We evaluate the function  $f$  at this point:

$$f(\lambda x_1 + (1 - \lambda)x_2, \lambda G_1 + (1 - \lambda)G_2) = \tilde{u}_i(\lambda x_1 + (1 - \lambda)x_2, \sigma(\lambda G_1 + (1 - \lambda)G_2)).$$

Without loss of generality, assume  $G_1 \leq G_2$ . Since  $\sigma$  is an increasing function, and the convex combination of  $G_1$  and  $G_2$  lies within the interval  $[G_1, G_2]$ , it follows that

$$\sigma(\lambda G_1 + (1 - \lambda)G_2) \in [\sigma(G_1), \sigma(G_2)].$$

Now, using the quasi-concavity of  $\tilde{u}_i$ :

$$\tilde{u}_i(\lambda x_1 + (1 - \lambda)x_2, \sigma(\lambda G_1 + (1 - \lambda)G_2)) \geq \min\{\tilde{u}_i(x_1, \sigma(G_1)), \tilde{u}_i(x_2, \sigma(G_2))\}.$$

This is precisely

$$f(\lambda x_1 + (1 - \lambda)x_2, \lambda G_1 + (1 - \lambda)G_2) \geq \min\{f(x_1, G_1), f(x_2, G_2)\}.$$

Thus, we have shown that  $f(x, G) = \tilde{u}_i(x, \sigma(G))$  is quasi-concave.

To determine whether goods  $x$  and  $G$  remain normal goods under the preferences of the composite utility function  $f(x, G) = \tilde{u}_i(x, \sigma(G))$ , we need to analyse how the demands for  $x$  and  $G$  change in response to a change in income:

Since  $\tilde{u}_i$  is a utility function in which  $x$  and  $G$  are normal goods, we know that  $\frac{\partial x}{\partial I} \Big|_{u_i} > 0$  and  $\frac{\partial G}{\partial I} \Big|_{u_i} > 0$ , where “ $|_{u_i}$ ” denotes *under the preferences of  $u_i$*  and  $I$  is income.

- Given that  $G$  is a normal good under the preferences of  $\tilde{u}_i$ , we know that  $\frac{\partial G}{\partial I} \Big|_{u_i} > 0$ , indicating that an increase in income leads to an increase in the demand for  $G$  under  $\tilde{u}_i$ .
- Now, consider the composite function  $f(x, G) = \tilde{u}_i(x, \sigma(G))$ . To demonstrate that  $G$  remains a normal good under the preferences of  $f$ , we need to show that  $\frac{\partial G}{\partial I} \Big|_f > 0$  also under these preferences.

Since  $G$  is an argument of  $\tilde{u}_i$ , we can use the chain rule to derive  $G$  with respect to  $I$  under the preferences of  $f$ :

$$\frac{\partial G}{\partial I} \Big|_f = \frac{\partial \tilde{u}_i}{\partial G} \cdot \frac{d\sigma}{dG} \cdot \frac{\partial G}{\partial I} \Big|_{u_i}$$

Given that  $\frac{\partial \tilde{u}_i}{\partial G}$  is positive (as  $G$  is a normal good under the preferences of  $\tilde{u}_i$ ), and  $\frac{d\sigma}{dG}$  is positive (as  $\sigma$  is an increasing function), we can see that  $\frac{\partial G}{\partial I} \Big|_f$  remains positive under the preferences of  $f$ . Hence, both goods  $x$  and  $G$  remain normal under the new utility function  $f(x, G) = \tilde{u}_i(x, \sigma(G))$  QED.

## Appendix B. Proof of Lemma 5.1 and Proposition 5.1

In a central lemma in [36], the authors go on to show the circumstances under which given a general public good game with a concave CES aggregator, a well-defined replacement function exists:

**Lemma B.1.** Let  $\Gamma = (\mathcal{A}_i, \tilde{u}_i)_{i \in I}$  be a generalised aggregative game featuring a concave CES social composition function. Under increasing, normal preferences, there is a unique  $g_i \in (0, w_i]$  for each player satisfying the first-order condition:

$$\frac{\partial \tilde{u}_i}{\partial g_i} \Big|_{\mathbf{g}_{-i}} \geq 0$$

with equality if  $g_i < w_i$ . Furthermore, the feasibility condition  $G^\rho \geq \beta_i g_i^\rho$  is satisfied for  $G \geq \bar{G}_i$  if  $0 < \rho < 1$  and for  $G \leq \bar{G}_i$  if  $\rho < 0$ .

This lemma implies that in general public good games involving a concave aggregator, the replacement correspondence  $r_i(G)$  is a well-defined function satisfying the first order conditions and has domain  $[\bar{G}_i, \infty)$  if  $0 < \rho < 1$  and  $[0, \bar{G}_i]$  if  $\rho < 0$ . A later proposition in that same paper clarifies the existence and uniqueness of equilibria in this model:

**Proposition B.1.** *Under the conditions of the Lemma above, the game has a unique equilibrium whenever the utility function  $\tilde{u}_i(x_i, G)$  is such that the corresponding indifference map is asymptotic to the axes.*

In a teamwork game, player  $i$ 's utility is a function of  $x_i$  and of the assessment  $\sigma(G)$ :

$$\tilde{u}_i(x_i, \sigma(G))$$

The properties of  $\sigma(G)$  in Definition 4.5 ensure that if  $i$ 's preferences represented by the utility function  $\tilde{u}_i(x_i, G)$  were increasing and normal, then they would remain increasing and normal under  $\tilde{u}_i(x_i, \sigma(G))$  as justified in Lemma 4.1. Hence, the Lemma in this appendix still applies in a teamwork game  $(\Gamma, \mathcal{T}, \sigma)$  whenever the aggregator (i.e. the teamwork outcome) is concave. This is the case of teamwork games involving a conjunctive task. In our case,  $g_i < w_i$  is replaced by  $g_i < p_i^L \cdot \Delta_i$ . Likewise, we can adapt the above proposition by using the teamwork game utility function  $\tilde{u}_i(x_i, \sigma(G))$ .

### Appendix C. Analysis of a player's replacement function in teamwork games with conjunctive tasks

Let us analyse how the replacement function in a conjunctive task ( $\rho < 1$ ) is impacted by changes in each of the elements of a teamwork game:

$$r_i(G)^{\rho-1} \left( \Delta_i - \frac{r_i(G)}{p_i^T} \right) = \frac{\alpha}{\beta_i p_i^T} \cdot G^{\rho-1} \cdot \frac{\sigma(G)}{\sigma'(G)}$$

- Deriving both sides of the equation w.r.t.  $G$  shows that the derivative of the left-hand side must be negative:

$$\frac{dr_i}{dG} \cdot r_i^{\rho-2} \cdot \left[ \rho \left( \Delta_i - \frac{r_i}{p_i^T} \right) - \Delta_i \right] = \frac{\alpha}{\beta_i p_i^T} G^{\rho-2} \left[ (\rho-1) \frac{\sigma(G)}{\sigma'(G)} + G \cdot \frac{\sigma'(G)^2 - \sigma''(G)\sigma(G)}{\sigma'(G)^2} \right]$$

Indeed, using that  $\frac{r_i(G)}{p_i^T} \in [0, \Delta_i]$  and  $(\rho-1) < 0$ , we see that the term in brackets in the left-hand side of the equation above must be negative. Thus, there are two behavioural regimes for  $r_i(G)$  as  $G$  increases

- If the term in brackets of the right-hand side of the equation is positive, then  $\frac{dr_i(G)}{dG} < 0$ . This is true whenever  $\frac{|\rho-1|}{G} < \frac{\sigma'(G)^2 - \sigma''(G)\sigma(G)}{\sigma'(G)\sigma(G)}$ .
- If the term in brackets of the right-hand side of the equation is negative,  $\frac{dr_i(G)}{dG} > 0$ , which happens whenever  $\frac{|\rho-1|}{G} > \frac{\sigma'(G)^2 - \sigma''(G)\sigma(G)}{\sigma'(G)\sigma(G)}$ .

This resonates with the qualitative regimes that Cornes and Hartley [36] distinguished (see Section 5.1.2): if  $\rho$  is sufficiently negative as in the second case, a substitution effect occurs and encourages player  $i$  to contribute more since a given increase in  $G$  generated by their contribution now has a lower opportunity cost in terms of private consumption. Otherwise, an income effect dominates and the preferred contributions by  $i$  will decrease as  $G$  increases.

- When we derive w.r.t.  $\alpha$ , we obtain the equation:

$$\frac{dr_i}{d\alpha} r_i^{\rho-2} \left( (\rho-1)\Delta_i - \frac{\rho}{p_i^T} r_i \right) = \frac{G^{\rho-1}}{\beta_i p_i^T} \cdot \frac{\sigma(G)}{\sigma'(G)}$$

The right-hand side of this equation is positive because  $\sigma'(G) > 0$ , and the term in parenthesis on the left is negative by the same argument as above. Hence, we conclude that  $\frac{dr_i(G)}{d\alpha} < 0$ . This result is in line with what was observed in Section 5.3.1: as the importance that a player gives to their leisure time increases, their optimal contributions decrease.

- Deriving w.r.t.  $\beta$ :

$$\frac{dr_i}{d\beta_i} r_i^{\rho-2} \left( (\rho-1)\Delta_i - \frac{\rho}{p_i^T} r_i \right) = -\frac{G^{\rho-1} \alpha}{\beta_i^2 p_i^T} \cdot \frac{\sigma(G)}{\sigma'(G)}$$

we see that the right-hand term is negative. Again, the factor in parenthesis on the left is negative. Hence, the derivative is positive to maintain the sign:  $\frac{dr_i(G)}{d\beta_i} > 0$ . The higher the weight or importance that a player's contribution has, the higher the optimal contribution will be.

- Regarding the term  $\frac{\sigma(G)}{\sigma'(G)}$ , we had already established that this coefficient represents the ratio between the incentive  $\sigma(G)$  and the slope of the tangent curve to the incentive  $\sigma(G)$ , and thus is a measure of *opportunity cost* in terms of leisure. Deriving with respect to it yields:

$$\frac{dr_i}{d(\sigma/\sigma')} r_i^{\rho-2} \left( (\rho-1)\Delta_t - \frac{\rho}{p_i^\tau} r_i \right) = \frac{\alpha}{\beta_i p_i^\tau} G^{\rho-1}$$

Using similar arguments to the above, it is straightforward to see that  $\frac{dr_i(G)}{d(\sigma/\sigma')} < 0$ . Just as in Section 5.3.1, higher values of this ratio (i.e. higher opportunity cost in terms of enjoyed leisure) induce lower contributions.

- When deriving with respect to the turn duration,  $\Delta_t$ , one obtains:

$$\frac{dr_i}{d\Delta_t} = \frac{p_i^\tau \cdot r_i}{\rho \cdot r_i + p_i^\tau (1-\rho)\Delta_t}$$

Using  $\rho - 1 < 0$ , we see that the fraction's numerator and denominator are positive. Thus, we conclude that  $\frac{dr_i(G)}{d\Delta_t} > 0$ , meaning that the more time a player has to perform a team task, the more they will contribute.

- Finally, the derivative w.r.t. the expertise level  $p_i^\tau$  yields, after some re-arrangement of the terms:

$$\frac{dr_i}{dp_i^\tau} \cdot \left[ r_i^{\rho-2} \cdot (\rho-1) \cdot \left( \Delta_t - \frac{r_i}{p_i^\tau} \right) - \frac{r_i^{\rho-1}}{p_i^\tau} \right] = -\frac{\alpha}{\beta_i (p_i^\tau)^2} G^{\rho-1} \frac{\sigma(G)}{\sigma'(G)} - \frac{r_i^\rho}{(p_i^\tau)^2}$$

Using that  $\rho - 1 < 0$  and  $r_i(G) \in [0, p_i^\tau \Delta_t]$ , we can see that  $\frac{dr_i(G)}{dp_i^\tau} > 0$ . Again, just as in the canonical case, higher expertise levels induce higher contributions to player  $i$ .

#### Appendix D. Proof of Lemma 5.2

Given a **general public good game** with Cobb-Douglas preferences (Equation (26)) and  $\rho > 1$ , Cornes and Hartley [36] demonstrate in a lemma the existence of three potential behavioural regimes within the set of best responses:

**Lemma D.1.** *There exists a threshold value  $G_{-i}^* > 0$  and a positive real-valued function  $\tilde{b}_i$  on  $[0, G_{-i}^*]$  such that:*

$$B_i(G_{-i}) = \begin{cases} \{0\} & \text{if } G_{-i} > G_{-i}^* \\ \{0, \tilde{b}_i(G_{-i})\} & \text{if } G_{-i} = G_{-i}^* \\ \{\tilde{b}_i(G_{-i})\} & \text{if } G_{-i} < G_{-i}^* \end{cases} \quad (\text{D.1})$$

Their proof revolves around the observation that the stationary points of the Cobb-Douglas utility function satisfy an expression that can be decomposed into a term dependent on  $g_i$ , denoted  $\phi(g_i)$ , and another dependent on  $G_{-i}$ :

$$\frac{d\tilde{u}_i}{dg_i} = \phi(g_i) - \alpha \cdot G_{-i}^\rho = 0 \quad (\text{D.2})$$

The authors establish that  $\phi(g_i)$  attains a maximum  $\phi_{\max}$ , implying that Equation (D.2) holds as long as  $\alpha \cdot G_{-i}^\rho$  is less than or equal to  $\phi_{\max}$ . Therefore,  $\phi_{\max} \geq \alpha \cdot G_{-i}^\rho$  serves as a necessary and sufficient condition for the stationary points of the utility function. Completing the proof involves: 1) demonstrating the well-defined nature of  $G_{-i}^*$ , 2) assigning the curve in the upper panel of Fig. 1 to the case  $G_{-i} > G_{-i}^*$ , 3) associating the middle panel with  $G_{-i} = G_{-i}^*$ , and 4) assigning the lower panel to the case  $G_{-i} < G_{-i}^*$ .

The approach that we will follow regarding teamwork games mirrors that of Cornes and Hartley: we will take our utility function to be

$$\tilde{u}_i = x_i^\alpha \cdot \sigma(G)$$

and we will break down  $\frac{\partial \tilde{u}_i}{\partial g_i}$  into two components: one dependent on  $(g_i, \sigma(G))$ , and another related to  $G_{-i}$ ; and we will prove that  $\phi(g_i, \sigma(G))$  attains a maximum. After that, we will show that  $G_{-i}^*$  is well-defined. Then, we will reason how each regime of  $G_{-i}$  relates to the panels in Fig. 1.

The stationary points of utility function satisfy:

$$\frac{\partial \tilde{u}_i}{\partial g_i} = \phi(g_i, \sigma(G)) - f(G_{-i})$$

By definition, the threshold points  $(g_i^*, G_{-i}^*)$  would satisfy  $\frac{\partial \tilde{u}_i}{\partial g_i} = 0$ . If we demonstrate that  $\phi(g_i, \sigma(G))$  has a maximum  $\phi_{\max} \forall G$ , then there will exist critical points  $(g_i^*, G_{-i}^*)$  as long as  $f(G_{-i}) \leq \phi_{\max}$ .

We have established that the condition  $\frac{\partial \tilde{u}_i}{\partial g_i} = 0$  is equivalent to:

$$\frac{\sigma(G)}{\sigma'(G)} \cdot G^{\rho-1} = \left( \Delta_t - \frac{g_i}{p_i^\tau} \right) \cdot \frac{\beta_i p_i^\tau}{\alpha} \cdot g_i^{\rho-1}$$

Manipulating this equation, we obtain:

$$\alpha^{\frac{\rho}{\rho-1}} (\beta_i g_i^\rho + G_{-i}^\rho) = \left[ \left( \Delta_t - \frac{g_i}{p_i^\tau} \right) \cdot \left( \frac{\sigma'(G)}{\sigma(G)} \right) \cdot \beta_i p_i^\tau \right]^{\frac{\rho}{\rho-1}} g_i^\rho$$

So we can write:

$$\phi(g_i, \sigma(G)) = \left[ \left( \Delta_t - \frac{g_i}{p_i^\tau} \right) \cdot \left( \frac{\sigma'(G)}{\sigma(G)} \right) \cdot \beta_i p_i^\tau \right]^{\frac{\rho}{\rho-1}} g_i^\rho - \alpha^{\frac{\rho}{\rho-1}} \beta_i g_i^\rho$$

Thus,  $\phi(0, \sigma) = 0$  and  $\phi(p_i^\tau \Delta_t, \sigma) = -\alpha^{\frac{\rho}{\rho-1}} \beta_i (p_i^\tau \Delta_t)^\rho$ . Fixing  $G$ , it's straightforward to verify that  $\phi(g_i, \sigma)$  has another root, that we denote  $c$ , at

$$c = p_i^\tau \Delta_t - \frac{\alpha}{\beta_i^{1/\rho}} \cdot \frac{\sigma(G)}{\sigma'(G)}$$

By Rolle's Theorem, since  $\phi$  is continuous for any  $g_i \in [0, p_i^\tau \Delta_t]$  and differentiable for any  $g_i \in (0, p_i^\tau \Delta_t)$ ,  $\phi$  will have at least one critical point for a fixed value of  $G$ . By the Weierstrass' Theorem, we know that the absolute maximum and minimum of  $\phi$  are attained on  $[0, p_i^\tau \Delta_t]$ . If we can demonstrate that at some critical point  $g'_i$ ,  $\phi(g'_i, \sigma(G)) > 0 \forall G$ , then there will be an absolute maximum in  $(0, p_i^\tau \Delta_t)$ .

Manipulating, we get the following equation for the critical points of  $\phi$  by imposing  $\frac{\partial \phi}{\partial g_i} = 0$ :

$$\left( \Delta_t - \frac{g_i}{p_i^\tau} \right)^k \cdot \left( \frac{\sigma'}{\sigma} \right)^k \cdot (\beta_i \cdot p_i^\tau)^k g_i^\rho \cdot \left( \frac{-k/p_i^\tau}{\Delta_t - \frac{g_i}{p_i^\tau}} + k \frac{\sigma(G)}{\sigma'(G)} \beta_i \left( \frac{g_i}{G} \right)^{\rho-1} \cdot \frac{\sigma \sigma'' - (\sigma')^2}{(\sigma')^2} + \frac{\rho}{g_i} \right) - \alpha^k \beta_i \rho g_i^{\rho-1} = 0$$

where  $k = \frac{\rho}{\rho-1}$ . We can extract the common factor  $\left( \Delta_t - \frac{g_i}{p_i^\tau} \right)^k \cdot \left( \frac{\sigma'(G)}{\sigma(G)} \right)^k \cdot (\beta_i \cdot p_i^\tau)^k g_i^\rho$  from this equation and substitute it into the expression of  $\phi(g_i, \sigma)$  to obtain the sign of  $g_i(c, \sigma)$ . Doing so, we verify that  $g_i(c, \sigma(G)) > 0 \forall G$  if and only if:

$$\frac{-k/p_i^\tau}{\Delta_t - \frac{g_i}{p_i^\tau}} + k \frac{\sigma \sigma'' - (\sigma')^2}{\sigma \cdot \sigma'} \beta_i \cdot \left( \frac{g_i}{G} \right)^{\rho-1} < 0$$

Equation (22) in Definition 4.5 of the evaluation function guarantees that this inequality holds for any value of  $G$ . Hence, the function  $\phi$  has a maximum  $\phi_{\max} \forall G$ , Q.E.D.

The correspondence of each panel in Fig. 1 to the different regimes of  $G_i$  follows the same reasoning as in [36]. Due to space constraints, we do not repeat that reasoning here.

We can solve for  $G_i^\rho$  in the global maximum condition (Eq. (36)) and substitute in the local maximum condition (35) to check if  $G_{-i}^*$  is well-defined:

$$\left( \Delta_t - \frac{g_i^*}{p_i^\tau} \right)^\alpha \cdot \sigma \left( \left( \frac{\sigma'(G_i^*)}{\sigma(G_i^*)} \frac{\beta_i}{\alpha} p_i^\tau \right)^{\frac{1}{\rho-1}} \cdot \left( \Delta_t - \frac{g_i^*}{p_i^\tau} \right)^{\frac{1}{\rho-1}} g_i^* \right) = \Delta_t^\alpha \cdot \sigma \left( \left( \left( \frac{\sigma'(G_i^*)}{\sigma(G_i^*)} \frac{\beta_i}{\alpha} p_i^\tau \right)^{\frac{\rho}{\rho-1}} \cdot \left( \Delta_t - \frac{g_i^*}{p_i^\tau} \right)^{\frac{\rho}{\rho-1}} - \beta_i \right)^{\frac{1}{\rho}} g_i^* \right)$$

For  $\beta_i > 0$ , the argument of  $\sigma$  in this equation's right-hand side (RHS) is smaller than that of the left-hand side (LHS). Given that  $\sigma$  is monotonically increasing, we have that the quotient  $\frac{\sigma_{\text{RHS}}}{\sigma_{\text{LHS}}} < 1$ . Thus, we can write

$$g_i^* = p_i^\tau \Delta_t \left( 1 - \left( \frac{\sigma_{\text{RHS}}}{\sigma_{\text{LHS}}} \right)^{1/\alpha} \right) > 0 \quad (\text{D.3})$$

By differentiating both sides of the global maximum condition (36) and substituting in the local maximum equation (35), we derive an expression of  $G_{-i}^*$  in terms of  $g_i^*$ :

$$\frac{\sigma(G_{-i}^*)}{\sigma'(G_{-i}^*)} G_{-i}^{*(\rho-1)} = \left( \Delta_t - \frac{g_i^*}{p_i^\tau} \right) \frac{\beta_i p_i^\tau}{\alpha} g_i^{*(\rho-1)}$$

which proves that  $G_{-i}^* > 0$ , confirming that  $G_{-i}^*$  is well-defined.

## Appendix E. Proof of Proposition 5.2

By Definition 3.5, the share correspondence is a function whenever the replacement correspondence is a function. In particular, given a teamwork game involving a disjunctive task and under the conditions of the Proposition, we know by Corollary 5.1 that both the null component and the positive component of the replacement correspondence are single-valued, and thus  $s_i(G)$  can also be decoupled in two components: a null one and a positive one defined over the same domains as  $r_i(G)$ . By definition of the share function, of  $G_i^*$  and of  $\bar{G}_i$ , we obtain the conditions that  $s_i(G)$  must satisfy in the proposition.

To see that the positive component of  $s_i$  is an increasing function, we use Eq. (35) and write it in terms of the share function (Definition 3.5):

$$G = \Delta_i \frac{p_i^T \cdot \beta_i^{1/\rho}}{s_i(G)^{1/\rho}} - \frac{\sigma(G)}{\sigma'(G)} \frac{\alpha}{s_i(G)} \quad (\text{E.1})$$

Which yields:

$$\frac{dG}{ds_i(G)} = \left( 1 + \frac{\sigma'(G)^2 - \sigma(G) \cdot \sigma''(G)}{\sigma'(G)^2} \cdot \frac{\alpha}{s_i(G)} \right)^{-1} \cdot \frac{1}{s_i(G)} \left( \frac{\sigma(G)}{\sigma'(G)} \frac{\alpha}{s_i(G)} - \frac{p_i^T}{\rho} \frac{\beta_i^{1/\rho} \Delta_i}{s_i(G)^{1/\rho}} \right)$$

Given the properties of the evaluation function, it is straightforward to see that the first factor in this equation is positive for all  $G$ . Using the condition  $\frac{\rho}{\Delta_i p_i^T} \geq \frac{\sigma'(\bar{G}_i)}{\sigma(\bar{G}_i)} \cdot \frac{\beta_i^{1/\rho}}{\alpha}$ , and the fact that for disjunctive tasks  $s_i(G)^{-1/\rho} < s_i(G)^{-1}$ , we can check that the second factor is positive too. Hence,  $s_i(G)$  is a strictly increasing function of  $G$  in the domain  $[G_i^*, \bar{G}_i]$ . Furthermore,  $s_i(G)$  is the inverse of a continuous function (Eq. (E.1)) on a compact domain and therefore itself continuous. These observations allow us to adapt Proposition 6.1 in [36] into Proposition 5.2 QED.

## Appendix F. Individual efforts in disjunctive and conjunctive tasks: tables

**Table F.10**

Individual effort across different teams for a **conjunctive** task ( $\rho = -10$ ) under three different assessment treatments. For simplicity, we fill only the upper diagonal of the matrices, which are symmetric.

(a) Soft Assessment Treatment ( $b = 3$ ).

$p_2 \backslash p_1$	0.30	0.50	0.70	0.90
0.30	(28%, 28%)	(46%, 30%)	(53%, 26%)	(56%, 22%)
0.50		(48%, 48%)	(53%, 40%)	(55%, 34%)
0.70			(44%, 44%)	(46%, 37%)
0.90				(38%, 38%)

(b) Medium Assessment Treatment ( $b = 5$ ).

$p_2^T \backslash p_1^T$	0.30	0.50	0.70	0.90
0.30	(29%, 29%)	(49%, 31%)	(56%, 27%)	(59%, 23%)
0.50		(57%, 57%)	(68%, 52%)	(72%, 45%)
0.70			(64%, 64%)	(67%, 55%)
0.90				(58%, 58%)

(c) Hard Assessment Treatment ( $b = 7$ ).

$p_2^T \backslash p_1^T$	0.30	0.50	0.70	0.90
0.30	(29%, 29%)	(49%, 31%)	(56%, 27%)	(59%, 23%)
0.50		(57%, 57%)	(69%, 52%)	(73%, 46%)
0.70			(69%, 69%)	(77%, 64%)
0.90				(72%, 72%)



**Table F.11**

Individual effort across different teams for a **disjunctive** task ( $\rho = 10$ ) under three different assessment treatments. For simplicity, we fill only the upper diagonal of the matrices, which are symmetric.

(a) Soft Assessment Treatment ( $b = 3$ ).

$p_2^T \backslash p_1^T$	0.30	0.50	0.70	0.90
0.30	[(63%, 0%), (0%, 63%)]	[(63%, 0%), (0%, 60%)]	(0%, 50%)	(0%, 41%)
0.50		[(60%, 0%), (0%, 60%)]	[(60%, 0%), (0%, 50%)]	[(60%, 0%), (0%, 41%)]
0.70			[(50%, 0%), (0%, 50%)]	[(50%, 0%), (0%, 41%)]
0.90				[(41%, 0%), (0%, 41%)]

(b) Medium Assessment Treatment ( $b = 5$ ).

$p_2^T \backslash p_1^T$	0.30	0.50	0.70	0.90
0.30	[(67%, 0%), (0%, 67%)]	[(67%, 0%), (0%, 78%)]	(0%, 71%)	(0%, 61%)
0.50		[(78%, 0%), (0%, 78%)]	(0%, 71%)	(0%, 61%)
0.70			[(71%, 0%), (0%, 71%)]	[(71%, 0%), (0%, 61%)]
0.90				[(61%, 0%), (0%, 61%)]

(c) Hard Assessment Treatment ( $b = 7$ ).

$p_2^T \backslash p_1^T$	0.30	0.50	0.70	0.90
0.30	[(67%, 0%), (0%, 67%)]	[(67%, 0%), (0%, 80%)]	(0%, 84%)	(0%, 78%)
0.50		[(80%, 0%), (0%, 80%)]	(0%, 84%)	(0%, 78%)
0.70			[(84%, 0%), (0%, 84%)]	[(84%, 0%), (0%, 78%)]
0.90				[(78%, 0%), (0%, 78%)]

## Data availability

Data will be made available on request.

## References

- [1] C.L. Apicella, J.B. Silk, *The evolution of human cooperation*, *Curr. Biol.* 29 (2019) R447–R450.
- [2] B. Dunin-Keplicz, R. Verbrugge, Teamwork in multi-agent systems: a formal approach, teamwork in multi-agent systems: a formal approach, <https://doi.org/10.1002/9780470665237>, 2010.
- [3] T. A. H. D. of the English Language (Fifth Edition), Teamwork, 2022.
- [4] S. Bowles, H. Gintis, *The Origins of Human Cooperation*, 2003.
- [5] U. N. D. of Economic, S. Affairs, Driving collective action for the sdgs: the role of further education - building a sustainable and resilient recover. Available at <http://tinyurl.com/unDrivingCollectiveAction>, 2021.
- [6] J. Von Neumann, O. Morgenstern, *Theory of Games and Economic Behavior*, 2nd rev., 1947.
- [7] E. Elkind, J. Rothe, Cooperative game theory, in: *Economics and Computation: an Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, 2016, pp. 135–193.
- [8] L.S. Shapley, A value for n-person games, in: *Contribution to the Theory of Games*, vol. 2, 1953.
- [9] R. Branzel, D. Dimitrov, S. Tijs, *Models in Cooperative Game Theory*, vol. 556, Springer Science & Business Media, 2008.
- [10] M. Maschler, S. Zamir, E. Solan, *Game Theory*, Cambridge University Press, 2020.
- [11] G. Chalkiadakis, E. Elkind, M. Wooldridge, *Computational Aspects of Cooperative Game Theory*, Springer Nature, 2022.
- [12] M. Tambe, Towards flexible teamwork, *J. Artif. Intell. Res.* 7 (1997) 83–124.
- [13] M. Wooldridge, *An Introduction to Multiagent Systems*, John Wiley & Sons, 2009.
- [14] E. Andrejczuk, J.A. Rodríguez-Aguilar, C. Sierra, A concise review on multiagent teams: contributions and research opportunities, in: *European Conference on Multi-Agent Systems*, Springer, 2016, pp. 31–39.
- [15] E. Andrejczuk, R. Berger, J.A. Rodríguez-Aguilar, C. Sierra, V. Marín-Puchades, The composition and formation of effective teams: computer science meets organizational psychology, *Knowl. Eng. Rev.* 33 (2018) e17.
- [16] D.S. Leslie, E.J. Collins, Individual q-learning in normal form games, *SIAM J. Control Optim.* 44 (2005) 495–514, <https://doi.org/10.1137/S0363012903437976>.
- [17] J. Jara-Ettinger, Theory of mind as inverse reinforcement learning, *Curr. Opin. Behav. Sci.* 29 (2019) 105–110.
- [18] B. Baker, Emergent reciprocity and team formation from randomized uncertain social preferences, *Adv. Neural Inf. Process. Syst.* 33 (2020) 15786–15799.
- [19] K. Ndousse, D. Eck, S. Levine, N. Jaques, Emergent social learning via multi-agent reinforcement learning, in: *International Conference on Machine Learning (ICML)*, 2021.
- [20] S.V. Albrecht, F. Christianos, L. Schäfer, *Multi-Agent Reinforcement Learning: Foundations and Modern Approaches*, MIT Press, 2024, <https://www.marl-book.com>.
- [21] K. Zhang, Z. Yang, T. Basar, Multi-agent reinforcement learning: a selective overview of theories and algorithms, in: *Handbook of Reinforcement Learning and Control*, vol. 18, 2021, p. 321.
- [22] J.Z. Leibo, V. Zambaldi, M. Lanctot, J. Marecki, T. Graepel, Multi-agent reinforcement learning in sequential social dilemmas, preprint, arXiv:1702.03037, 2017.
- [23] E. Hughes, J.Z. Leibo, M. Phillips, K. Tuyls, E. Dueñez-Guzmán, A. García Castañeda, I. Dunning, T. Zhu, K. McKee, R. Koster, et al., Inequity aversion improves cooperation in intertemporal social dilemmas, *Adv. Neural Inf. Process. Syst.* 31 (2018).
- [24] N. Jaques, A. Lazaridou, E. Hughes, C. Gulcehre, P. Ortega, D. Strouse, J.Z. Leibo, N. De Freitas, Social influence as intrinsic motivation for multi-agent deep reinforcement learning, in: *International Conference on Machine Learning*, PMLR, 2019, pp. 3040–3049.
- [25] K.R. McKee, I. Gemp, B. McWilliams, E.A. Duñez-Guzmán, E. Hughes, J.Z. Leibo, Social diversity and social preferences in mixed-motive reinforcement learning, preprint, arXiv:2002.02325, 2020.

- [26] M. Rodríguez-Soto, M. Lopez-Sanchez, J.A. Rodríguez Aguilar, Multi-objective reinforcement learning for designing ethical environments, in: Z.-H. Zhou (Ed.), *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21, International Joint Conferences on Artificial Intelligence Organization*, 2021, pp. 545–551, main Track.
- [27] U. Madhushani, K.R. McKee, J.P. Agapiou, J.Z. Leibo, R. Everett, T. Anthony, E. Hughes, K. Tuyls, E.A. Duñez-Guzmán, Heterogeneous social value orientation leads to meaningful diversity in sequential social dilemmas, preprint, arXiv:2305.00768, 2023.
- [28] J. Kramár, T. Eccles, I. Gemp, A. Tacchetti, K.R. McKee, M. Malinowski, T. Graepel, Y. Bachrach, Negotiation and honesty in artificial intelligence methods for the board game of diplomacy, *Nat. Commun.* 13 (2022) 7214.
- [29] A. Dafeo, E. Hughes, Y. Bachrach, T. Collins, K.R. McKee, J.Z. Leibo, K. Larson, T. Graepel, Open problems in cooperative ai, preprint, arXiv:2012.08630, 2020.
- [30] I. Gemp, K. McKee, R. Everett, E.A. Duenez-Guzman, Y. Bachrach, D. Balduzzi, A. Tacchetti, D3c: reducing the price of anarchy in multi-agent learning, <https://openreview.net/forum?id=8wa7HrUsELL>, 2021.
- [31] M.K. Jensen, et al., *Aggregative games*, in: *Handbook of Game Theory and Industrial Organization*, vol. 1, 2018, pp. 66–92.
- [32] K. Taywade, B. Harrison, A. Bagh, J. Goldsmith, Modelling cournot games as multi-agent multi-armed bandits, in: *The International FLAIRS Conference Proceedings*, vol. 35, 2022, <https://journals.flvc.org/FLAIRS/article/view/130697>.
- [33] Y. Qian, C. Zhang, B. Krishnamachari, M. Tambe, Restless poachers: handling exploration-exploitation tradeoffs in security domains, in: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, 2016, pp. 123–131.
- [34] K. Lewin, *Resolving Social Conflicts; Selected Papers on Group Dynamics*, 1948.
- [35] I.D. Steiner, *Group Process and Productivity*, Academic Press, New York, 1972.
- [36] R. Cornes, R. Hartley, Weak links, good shots and other public good games: building on bbv, *J. Public Econ.* 91 (2007) 1684–1707.
- [37] R. Cornes, R. Hartley, Aggregative public good games, *J. Public Econ. Theory* 9 (2007) 201–219.
- [38] L. Hernández-Lamonedá, F. Sánchez-Sánchez, Rankings and values for team games, *Int. J. Game Theory* 39 (2010) 319–350.
- [39] T. Hiller, Structure of teams—a cooperative game theory approach, *Manag. Decis. Econ.* 40 (2019) 520–525.
- [40] T. Yan, C. Kroer, A. Peysakhovich, Evaluating and rewarding teamwork using cooperative game abstractions, *Adv. Neural Inf. Process. Syst.* 33 (2020) 6925–6935.
- [41] D.P. Myatt, C. Wallace, Evolution, teamwork and collective action: production targets in the private provision of public goods, *Econ. J.* 119 (2009) 61–90.
- [42] J. Carpenter, S. Bowles, H. Gintis, S.-H. Hwang, Strong reciprocity and team production: theory and evidence, *J. Econ. Behav. Organ.* 71 (2009) 221–232.
- [43] A.A. Baños, P.H. Rojas, Monitoring teamwork: how do punishment motivations change with group size, *Econ. Indust.* (2019) 51–59.
- [44] T. Kameda, M.F. Stasson, J.H. Davis, C.D. Parks, S.K. Zimmerman, Social dilemmas, subgroups, and motivation loss in task-oriented groups: in search of an “optimal” team size in division of work, *Soc. Psychol. Q.* (1992) 47–56.
- [45] L.C. Corchón, *Aggregative games*, *SERIEs* 12 (2021) 49–71.
- [46] D. Acemoglu, M.K. Jensen, Aggregate comparative statics, *Games Econ. Behav.* 81 (2013) 27–49.
- [47] T.C. Bergstrom, L.E. Blume, H.R. Varian, Uniqueness of Nash Equilibrium in Private Provision of Public Goods: an Improved Proof, 1992.
- [48] J. Hirshleifer, From weakest-link to best-shot: the voluntary provision of public goods, *Public Choice* 41 (1983) 371–386.
- [49] R. Cornes, Dyke maintenance and other stories: some neglected types of public goods, *Q. J. Econ.* 108 (1993) 259–271.
- [50] R. Cornes, *Voluntary Contribution Model of Public Goods*, Palgrave Macmillan UK, London, 2016, pp. 1–9.
- [51] W. Buchholz, D. Rübhelke, *The Theory of Externalities and Public Goods*, Springer, 2017.
- [52] M.K. Jensen, Diversity and incentivized team performance, Unpublished. Available at <http://tinyurl.com/jensenDiversityPerformance>, 2022.
- [53] D. De Cremer, Trust and fear of exploitation in a public goods dilemma, *Curr. Psychol.* 18 (1999) 153–163.
- [54] D. Balliet, P.A. Van Lange, Trust, conflict, and cooperation: a meta-analysis, *Psychol. Bull.* 139 (2013) 1090.
- [55] J.E. McGrath, *Social Psychology: A Brief Introduction*, 1964.
- [56] M.A. Campion, G.J. Medsker, A.C. Higgs, Relations between work group characteristics and effectiveness: implications for designing effective work groups, *Pers. Psychol.* 46 (1993) 823–847.
- [57] N.L. Kerr, R.S. Tindale, Group performance and decision making, *Annu. Rev. Psychol.* 55 (2004) 623–655.
- [58] C. Casal, F. Alonso, Á. Jiménez, J. Domínguez, *PSICOLOGÍA DE LOS GRUPOS*, Colección Grado, UNED, <https://books.google.es/books?id=2R1vtwAACAAJ>, 2011.
- [59] A. Caparrós, M. Finus, The corona-pandemic: a game-theoretic perspective on regional and global governance, *Environ. Resour. Econ.* 76 (2020) 913–927.
- [60] N.L. Kerr, Motivation losses in small groups: a social dilemma analysis, *J. Pers. Soc. Psychol.* 45 (1983) 819.
- [61] S.J. Karau, K.D. Williams, Social loafing: a meta-analytic review and theoretical integration, *J. Pers. Soc. Psychol.* 65 (1993) 681.
- [62] R. Cornes, T. Sandler, *The Theory of Externalities, Public Goods, and Club Goods*, Cambridge University Press, 1996.
- [63] E. Fehr, K.M. Schmidt, A theory of fairness, competition, and cooperation, *Q. J. Econ.* 114 (1999) 817–868.
- [64] K.D. Williams Karau, S.J. Karau, Social loafing and social compensation: the effects of expectations of co-worker performance, *J. Pers. Soc. Psychol.* 61 (1991) 570–581.
- [65] T. Roughgarden, Algorithmic game theory, *Commun. ACM* 53 (2010) 78–86.
- [66] C.M. Wu, E. Schulz, M. Speekenbrink, J.D. Nelson, B. Meder, Generalization guides human exploration in vast decision spaces, *Nat. Hum. Behav.* 2 (2018) 915–924.
- [67] P.B. Reverdy, V. Srivastava, N.E. Leonard, Modeling human decision making in generalized gaussian multiarmed bandits, *Proc. IEEE* 102 (2014) 544–571.
- [68] R.S. Sutton, A.G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 2018.
- [69] A.E. Roth, I. Erev, Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term, *Games Econ. Behav.* 8 (1995) 164–212.
- [70] I. Erev, A.E. Roth, Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria, *Am. Econ. Rev.* (1998) 848–881.
- [71] C. Child, C. Koluman, T. Weyde, Modelling emotion based reward valuation with computational reinforcement learning, in: *Proceedings of the 41st Annual Conference of the Cognitive, Science Society*, 2019, pp. 582–588.
- [72] G. Gigerenzer, W. Gaissmaier, Heuristic decision making, *Annu. Rev. Psychol.* 62 (2011) 451–482.
- [73] R.D. McKelvey, T.R. Palfrey, Quantal response equilibria for normal form games, *Games Econ. Behav.* 10 (1995) 6–38.
- [74] D. McFadden, et al., Quantal choice analysis: a survey, *Ann. Econ. Soc. Meas.* 5 (1976) 363–390.
- [75] N. Tumennasan, To err is human: implementation in quantal response equilibria, *Games Econ. Behav.* 77 (2013) 138–152.
- [76] T. Kozitsina, I. Kozitsin, I. Menshikov, Quantal response equilibrium for the prisoner’s dilemma game in markov strategies, *Sci. Rep.* 12 (2022) 4482.
- [77] J.K. Goeree, C.A. Holt, T.R. Palfrey, *Quantal Response Equilibrium: A Stochastic Theory of Games*, Princeton University Press, Princeton, 2016.
- [78] G. Ge, G. Godager, Predicting strategic medical choices: an application of a quantal response equilibrium choice model, *J. Choice Model.* 39 (2021) 100282.
- [79] S.P. Anderson, J.K. Goeree, C.A. Holt, The logit equilibrium: a perspective on intuitive behavioral anomalies, *South. Econ. J.* 69 (2002) 21–47, <http://www.jstor.org/stable/1061555>.
- [80] S. Govindan, P.J. Reny, A.J. Robson, et al., A short proof of harsanyi’s purification theorem, *Games Econ. Behav.* 45 (2003) 369–374.
- [81] R. Hartley, R. Cornes, Joint Production Games and Share Functions, Keele Department of Economics Discussion Papers (1995-2001) 2000/07, Department of Economics, Keele University, 2000, <https://ideas.repec.org/p/kee/keeldp/2000-07.html>.

- [82] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: a next-generation hyperparameter optimization framework, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2019.
- [83] D.L. Feltz, N.L. Kerr, B.C. Irwin, Buddy up: the Köhler effect applied to health games, *J. Sport Exerc. Psychol.* 33 (2011) 506–526.
- [84] F. Christianos, G. Papoudakis, M.A. Rahman, S.V. Albrecht, Scaling multi-agent reinforcement learning with selective parameter sharing, in: International Conference on Machine Learning, PMLR, 2021, pp. 1989–1998.