Autonomous Robots (2019) 43:359–373
https://doi.org/10.1007/s10514-018-9767-9

“I Choose... YOU!” Membership preferences in human–robot teams

Filipa Correia[1] · Sofia Petisca[2] · Patrícia Alves-Oliveira[2] · Tiago Ribeiro[1] · Francisco S. Melo[1] · Ana Paiva[1]

Received: 7 December 2017 / Accepted: 22 May 2018 / Published online: 8 June 2018

Abstract

Although groups of robots are expected to interact with groups of humans in the near future, research related to teams of humans and robots is still scarce. This paper contributes to the study of human–robot teams by describing the development of two autonomous robotic partners and by investigating how humans choose robots to partner with in a multi-party game context. Our work concerns the successful development of two autonomous robots that are able to interact with a group of two humans in the execution of a task for social and entertainment purposes. The creation of these two characters was motivated by psychological research on learning goal theory, according to which we interpret and approach a given task differently depending on our learning goal. Thus, we developed two robotic characters implemented in two robots: *Emys* (a competitive robot, based on characteristics related to performance-orientation goals) and *Glin* (a relationship-driven robot, based on characteristics related to learning-orientation goals). In our study, a group of four (two humans and two autonomous robots) engaged in a card game for social and entertainment purposes. Our study yields several important conclusions regarding groups of humans and robots. (1) When a partner is chosen without previous partnering experience, people tend to prefer robots with relationship-driven characteristics as their partners compared with competitive robots. (2) After some partnering experience has been gained, the choice becomes less clear, and additional driving factors emerge as follows: (2a) participants with higher levels of competitiveness (personal characteristics) tend to prefer Emys, whereas those with lower levels prefer Glin, and (2b) the choice of which robot to partner with also depends on team performance, with the winning team being the preferred choice.

Keywords Social robots · Human–robot teams · Collaboration

1 Introduction

As robots become a more common and pervasive technology in our society, they will slowly become common aids in our everyday activities. For robots to co-exist with humans in their everyday activities, it is fundamental that they are able to respond to the multiple needs of humans in a wide range of domains and situations; that they are flexible to accommodate the varied and varying operation conditions imposed by such multi-purposefulness; and, perhaps more importantly, that they are able to successfully integrate explicit and implicit operational and social protocols in their place of actuation. In other words, it is imperative that the robots are able to integrate into teams with other robots and especially with humans. Therefore, it is inevitable that humans and robots will engage in cooperative activities as teams in the near future (Coradeschi and Saffiotti 2006). However, for this to be possible, robots must be endowed with the necessary competencies that enable social interaction with humans to occur.

The present paper is an extended version of the work in the article “Groups of humans and robots: understanding membership preferences and team formation”, published in the Proceedings of Robotics: Science and Systems (2017), with the https://doi.org/10.15607/RSS.2017.XIII.024. The present version includes a detailed description of the autonomous robots’ development, not included in the aforementioned article. It also includes a significantly improved discussion of our results in terms of human–robot collaboration.

This is one of several papers published in *Autonomous Robots* comprising the “Special Issue on Robotics Science and Systems”.

✉ Filipa Correia
filipacorreia@tecnico.ulisboa.pt

1 INESC-ID & Instituto Superior Técnico, Universidade de Lisboa, Avenida Professor Cavaco Silva, Edifício IST, 2744-016 Porto Salvo, Portugal

2 Instituto Universitário de Lisboa, ISCTE-IUL (CIS-IUL) & INESC-ID, Lisbon, Portugal

Several works have explored different ways of interacting in teams comprising both robots and humans (Chang et al. 2012; Eyssel and Kuchenbrandt 2012; Fraune et al. 2015). In this paper, we contribute to this literature, exploring the social competences necessary to build successful teams of humans and robots. In particular, we study multi-robot and multi-person interactions, investigating people's preferences toward robots with different goals and exhibiting different social behaviours.

In this context, several questions arise: What will a team consisting of humans and robots be like? Will people wilfully partner with a robot? If so, what type of robot will they prefer as a teammate? The latter question is particularly interesting, as the choice of whom to partner with on a team depends on many different factors, including the characteristics of the individuals and the tasks to be executed. This is also true for teams consisting solely of humans, as differences in the social competencies or personality of the other may greatly influence the degree to which a human is willing to partner with another (Hinds et al. 2000). According to Hinds et al. (2000), when we, as humans, select a team member to work with, we tend to privilege homogeneous groups with high indicators of competence and with greater similarity and familiarity. These preferences may be related to our attempts to make choices that will maximise our expectation of success. Similarly, previous research in (HRI) demonstrated that users prefer robots whose personalities match their own in terms of introversion/extroversion in a therapeutic task (Tapus et al. 2008). This paper extends these results, exploring the influence of robot traits in both cooperative and competitive interactions.

We investigate multi-robot multi-person interactions in the context of a game to study team formation and preferences. We describe an investigation of team formation with robots, considering robots with different *goal orientations*. It has been shown that, at an individual level, people's goal orientations have a major effect on how they approach and respond to a task. Dweck (1986) extended the notion of goal orientation introduced by Eison (1979) and concluded that during a task, people will present either a *learning goal* (i.e., an interest in learning something) or a *performance goal* (i.e., an interest in the result and what judgements will emerge from it). Teams consisting of individuals with a "learning-goal" orientation are reported to show high levels of mutual support behaviours and high qualities of interaction, team efficacy and commitment. By contrast, teams consisting of individuals with a "performance-goal" orientation are negatively correlated with team efficacy and commitment (Porter 2005).

In our work, we explore the two aforementioned goal orientations and develop two different robotic characters: (1) a more *relationship-driven* character (named Glin), mapping to the learning-goal orientation, and (2) a more *competitive* one (named Emys), mapping to the performance-goal orientation. These two robots autonomously play a card game with two humans companions. Thus, the autonomous robots interact in a group of four and are partnered with either a human or the other robot. Two studies are reported in this paper, and their study methodologies are illustrated in this link.[1]

The first study provided a **validation the goal orientations of both characters**. The second study investigated **which robotic character is preferred as a partner** in the game, depending on the two goal orientations. To this end, robots and humans were placed on teams and played a game in a social setting. Both robots played equally well, but one robot displayed very competitive interactions, whereas the other was more interested in the quality of the interactions. We expected that, (i) overall, participants tend to choose the more relational robot as their preferred partner and that (ii) the level of competitiveness played a role in the choice, with higher scores being associated with choosing the competitive robot. Our results showed that, in general, the participants liked having both robots as partners and that, upon first impressions, they do preferred the more relational robot to the competitive one. However, the results also showed that after repeated interactions and after partnering with both robots, the choices of the participants became less clear and other driving factors emerged in their decision. For example, more competitive people preferred the more competitive robot. Furthermore, our findings also showed that team performance affects the partner choice.

In general, this paper contributes to the study of teams of humans and autonomous robots; specifically, it addresses membership in mixed teams in the context of playing a game. Furthermore, this paper presents evidence that preferences regarding team formation with an autonomous robot depend not only on the robot's goal orientation (competitive vs. relationship-driven) but also on the characteristics of the people involved. This finding has implications for the field of (HRI), as it introduces factors that can impact preferences regarding the choice of a robotic partner and how these preferences vary over time.

2 Related work

Over the years, the field of (HRI) has evolved from being mostly focused on "one-to-one" interactions to considering more complex scenarios in which (1) *individual users* interact with *multiple robots* or (2) *multiple users* interact with *individual robots*. However, a vision for the future concerns not only individual robots operating in a variety of human

[1] Video explaining the methodology design of the two studies reported in the present paper: https://youtu.be/rwvBIDsN6Cc.

environments (Gates 2007) but also multiple robots interacting with multiple people and with each other (Rus et al. 1995). Indeed, Groom and Nass (2007) identified this trend by defining several benchmarks that enable the emergence of optimal social teams consisting of humans and robots and by emphasising how robots can complement and improve current human-human teams. Furthermore, well-established and grounded social psychological theories postulate that people's behaviour changes when they are exposed to a certain group or individual (e.g., Bornstein and Yaniv 1998), and therefore, the study of groups of humans and robots is undeniably a crucial area of (HRI).

For social robots to be able to interact with multiple users, they need to be endowed with social competencies. In general, the research findings suggest that humans often treat instruments of technology as social actors (Reeves and Nass 1996), applying social rules and expectations and exhibiting overlearned social behaviours, such as politeness towards machines (Nass and Moon 2000). Several studies have been performed to analyse group effects related to individual robots, such as group membership and social categorisation (e.g., Kuchenbrandt et al. 2013). In addition, studies on (HRI) have confirmed that social categorisation processes associated with groups also generalise to social robots. By manipulating group membership, Eyssel and Kuchenbrandt (2012) showed that people anthropomorphise and prefer an in-group robot to a greater extent than an out-group robot. Chang et al. (2012) studied the type of behaviour that humans choose to adopt (competitive or cooperative) depending on the group size (a group of humans or an individual human player). The results showed that participants behave more competitively towards a robot when they are in a group than when they are interacting as individual players. Additionally, a cross-cultural field study investigated participants' behaviour depending on the number of robots (a single robot or a group of robots) with which they were interacting and the type of behaviour (social or functional) the robot(s) exhibited. The results showed that people regarded single social robots more positively than social robots in a group. By contrast, people felt more positively towards a group of functional robots than towards a single functional robot. This research already suggests that the specific characteristics of robots, in this case, functional versus social behaviour, influence their group effects (Fraune et al. 2015). Moreover, personality appears to be an important variable influencing how people perceive and choose robots. Findings suggest that people tend to prefer robots whose personalities match their own in a therapeutic context (Tapus et al. 2008), with similar findings related to pet-like robots (Lee et al. 2006). Additionally, it has been found that people's personality traits are predictive of comfortable proximity distances when interacting with social robots (Walters et al. 2005). More recently, Fraune et al. (2017) explored people's responses to groups of robots and compared the responses to different types of groups by varying the "diversity" of the groups. A (WoZ) approach was used to control the robots, and the human participants were directed to solve a task in the presence of 3 robots (under two conditions: high similarity and diversity). The results of this experiment showed that people perceive multiple robots that act and look the same as more threatening than a diverse group of robots.

Regarding human–robot teams, other concerns arise when exploring how these partnerships can evolve in a symbiotic manner and contribute to improved human–robot collaboration. For example, the efficiency of work performed with a robot increases when the robot shares non-verbal cues with its teammate (Breazeal et al. 2005). Furthermore, Shah et al. (2011) have shown that team performance increases when the behaviour of the robot is based on human-human cooperation, and the same is seen when the robot adapts to the user (Hoffman and Breazeal 2007). Another study involving groups of humans and robots in a team examined the role of backchanneling competencies in a robot. The results of this study support the assumption that backchanneling is important for team performance even for robots (Jung et al. 2013). All of this brings attention to the numerous factors and characteristics that influence users when working with robots. To create better human–robot teams, these factors and characteristics should be considered to understand which types of robot characteristics are best for each specific context.

Despite the significant work that has been done in this new area of research, we believe that the work presented here makes a novel contribution to this new era, moving beyond "one-to-one" and "one-to-many" interactions to scenarios in which several robots and several humans are interacting with each other. Moreover, we also contribute to the field by demonstrating how some preferences regarding robotic partners are influenced by the social characteristics of both the humans and the robots. Finally, we contribute to the field of (HRI) through the development of two autonomous robots that can interact with each other and with two humans.

3 The card game scenario

To explore the topic of human–robot teams and the role of goal orientation in the formation of these teams, it was necessary to define a suitable scenario to study this phenomenon. We chose the card game Sueca, which is a four-player game played between two teams. Partners on the same team sit across from each other and must play symbiotically to succeed in winning the game against their opponents. Because this is a hidden-information card game in which players do not know each other's cards, the relationship between each player and his/her partner constitutes an especially relevant part of the game. Traditionally, two partners who frequently play with each other do not want to switch to different teams, as they have often developed communication signals or other complicit mechanisms and each partner understands how the other plays, thus making them a better team.

Players of SUECA typically do not start the game on equal footing, as it depends on the initial distribution of the cards that are dealt by the players at the beginning of each game. One team might have a higher probability of winning the game than the other, and one player might have more opportunities to make a good play than the others depending on the initial card distribution. Therefore, the environment of the game is considered inaccessible (Russell et al. 1995), which complicates the task for any autonomous agent. Moreover, in the case of a robotic agent, this task mirrors the real game experience that humans have during card games, making this a very natural and believable scenario or even more complex in comparison to scripted interactions or (WoZ)-built scenarios.

In our scenario, two robots will play with two human players in different game settings, i.e., we established different teams: robot-robot versus human-human and human–robot versus human–robot. We consider a mixed environment (see Fig. 1) in which humans play with physical cards and robots play with virtual cards. The human players hold their cards during the game and are responsible for shuffling them and distributing them to each player. To assist with the game play, the physical cards have fiducial markers that can be detected by a multi-touch surface, thereby perfectly blending the natural card game experience for the humans with the digital version required by the robots. As a result, the robots must autonomously play the game (with virtual cards), which is unpredictable for both the humans and the robots; consequently, this is a very realistic scenario.

4 Developing a robotic game player

Two aimed robotic characters were embedded in similar robotic agents. The current section describes the development process for a robotic agent that acts autonomously. It begins with an analysis of a user-centred study that further inspired the definition of the agent's perceptions as well as its behaviours, which together constitute the main perception-reaction loop of each robotic game player.

4.1 User-centred design

We conducted a user-centred study to analyse and collect the behaviours of human players during the game interaction, as described in Correia et al. (2016). The video records of 10 independent games were annotated and converted into utterances of verbal and non-verbal behaviours. In addition to the dialogues, these utterances allowed us to analyse the relevant gaze directions during this card game, such as gazing at the table where the game was being played, at the partner, at opponents, at their own hand (own cards), or elsewhere. The initial coding scheme was semi-structured according to the following game stages that are commonly known: *session start*, *shuffle deck*, *split deck*, *deal cards*, *game start*, *next player turn*, *player played*, *trick end*, *game end*, and *session end*. Additionally, we realised a discrepancy in the competitiveness levels of some players between their interactions towards partners and towards opponents, especially during the game stages of *next player turn*, *player played*, *trick end*, *game end*, and *session end*, which we classified as the ones eliciting the competitiveness of the players.

Finally, we further examined the content of behaviours triggered during the game stages of *player played* and *trick end*, which revealed more complex appraisals. The participants of our user-centred study usually expressed how desirable such events are regarding their current scores in the game. Nevertheless, they were careful not to disclose information about their cards, which are unknown for the other players due to the fact that SUECA is a hidden-information game.

This detailed behavioural analysis was crucial to creating a social robotic player that acts in a natural and human fashion. Moreover, it guided the development of the two proposed characters (detailed in Sect. 5), where one is more performance-oriented and competitive, i.e., Emys, while the other is more learning-oriented and relationship-driven, i.e., Glin.

4.2 Perceptions of the robot

After having analysed the relevant game stages that trigger the interactions among players during this card game, we were able to define the perceptions of each of the robotic players.

We decided to confine the perceptions of the robotic player solely to game events, without any further speech or image recognition. Therefore, a robotic game player requires the notification of all relevant game events, as well as their associated information, i.e., the identification of the player and/or team performing the event or the card being played. It is crucial that this communication occurs in run time so that the robotic player can autonomously react at the right moment.

Regarding the game events that elicit the competitiveness of a player, their perception process includes two additional appraisals. The first appraisal attempts to assess the immediate impact of a game event. It begins by identifying the team that benefits the most from it. In the case of a *trick end*, *game end*, or *session end*, this identification refers to the winning team. In the case of a *player played* event, it refers to the current trick winner. Although the ultimate winner of a trick can only be assessed after the four moves have been completed, human players usually acknowledge indefinite winners during the trick. Consequently, it is also important to assess the previous trick winner to check if the current move caused a change in the current scores. Finally, it is also relevant to quantify the impact of the play on the game scores. We have defined two boundaries based on the value of the highest cards of each suit in the SUECA game: *low* impact when the move/trick adds at least 3 points[2] and *high* impact when the move/trick adds at least 10 points.[3] The combination of these three factors is shown in Table 1 and is ordered according to how favourable the game event is at the moment for the agent's team.

The second appraisal process is an emotional appraisal related to the overall impact on the task. The agent uses an OCC model (Ortony et al. 1990) with the appraisal variables of *Desirability*, *Desirability For Others*, and *Goal Likelihood*. Each agent has the goal of "winning the set of games". According to the following equations, the first two appraisal variables are updated after each *player played* event, while the last one is updated after each *trick end* event.

$$Desirability = \frac{min(max(-15, TP), 15)}{15}$$
$$DesirabilityForOthers = -Desirability$$
$$G.Likelihood = 0,5 \times \frac{MySP}{MySP+OSP} + 0,5 \times \frac{MyGP}{MyGP+OGP}$$

$\|TP\|$ is the current trick points, and its valence may be positive if the current trick winner is the agent's team or negative if it is the other team. $MySP$ are the current session points of the agent's team, while OSP are the current session points of the other team. Similarly, $MyGP$ are the current game points of the agent's team, and OGP are the current game points of the other team. According to the OCC model, these appraisal variables will generate different types of emotions: well-being (i.e., joy and distress), fortune for others (i.e., resentment, gloating, happy for, and pity), and prospect-based (i.e., hope and fear).

The described appraisal mechanisms perform identically for our two robotic game players. Nevertheless, the role each player has on the game will produce different appraisals and, therefore, different perceptions. For instance, a player that adds more points to his/her team will be positively appraised by its robotic partner and negatively appraised by its robotic opponent.

4.3 Behaviours of the robot

The final step in developing a robotic game player and closing the reaction-action loop is to define its behaviours. The goals of each robotic player include both socially interacting during the game and efficiently performing the task of playing the game, which results in both social and task-related actions, described as follows.

4.3.1 Social behaviours

The social behaviours of a robot are limited to the features of its embodiment. For the purposes of this card game scenario, we chose the stationary robotic head (EMYS) (Kedzierski et al. 2013) capable of using gazes, animations, postures, and dialogues, which are expressed as follows.

The basic reaction to all the perceptions is a change in the gaze direction. For instance, game events such as *shuffle deck*, *split deck*, or *deal deck* cause an immediate shift in the gaze direction of the robots towards the player performing these actions. In the same manner, game events referring to a new card on the table lead to a gaze towards the table centre. For the *next player turn* events, the gaze instructions depend on whether the next player is the robot itself or another player. In the former case, the robot immediately gazes at its hand (own cards). In the last case, it waits 2 s before looking to the next player to simulate that it is acknowledging the previous play.

Regarding other non-verbal behaviours, the robotic game player produces expressive facial animations and physical postures according to its activated emotional state. This is the behavioural response of the emotional appraisal previously described, which aims to display the perception of its performance on the task. As different emotions can be simultaneously activated by a single event (e.g., sadness and hope), the strongest emotion among the activated emotional states is used to drive the robot's physical posture and to select some of its animations. These non-verbal emotional behaviours perform similarly on each robotic game player, as their emotional appraisals are similar.

As for the robotic agent's more complex reactive behaviours in the form of utterances (which include dialogue, animations and/or gaze instructions), they appear as a reaction to a particular game event. Moreover, reactions will differ according to different perceptions of the game events, as shown in Table 2. Based on the user-centred study, competitive behaviours usually occur after extremely favourable or unfavourable game events, which, according to the appraisal presented in Table 1, occur in the highest and lowest situations. The authoring of the utterances is detailed in Sect. 5, as they were used to convey the manipulation of the goal orientation on the robotic players. Nevertheless, it is important to mention that all the utterances begin with a gaze towards the player the robot is interacting with and may end with a gaze towards the next player, if applicable.

[2] Value of the 2nd lowest card.

[3] Value of the 2nd highest card.

Table 1 Appraisal of game events ordered by the benefit they add to the agent's current trick score

Previous trick winner	Current trick winner	Impact on score	Valence
Other team	My team	High	+
Other team	My team	Low	
My team	My team	High	
My team	My team	Low	
Other team	Other team	Low	–
Other team	Other team	High	
My team	Other team	Low	
My team	Other team	High	

4.3.2 Task-related behaviours

For each robotic player to succeed in its task of playing the card game, it contains an algorithm able to perform online computations within the SUECA domain and to choose a suitable card to play.

Considering the main property of this card game, i.e., the imperfect information, we considered implementing the (PIMC) algorithm. It is based on the Monte Carlo method, which has recently successfully solved similar games, such as Bridge (Ginsberg 2001), Skat (Buro et al. 2009), or Hearts (Sturtevant 2008).

Algorithm 1 PIMC search pseudo-code.

```
procedure PIMC(InfoSet I, int N)
    for all m ∈ Moves(I) do
        val[m] = 0
    for all i ∈ {1..N} do
        x = Sample(I)
        for all m ∈ Moves(I) do
            val[m] += PerfectInfoValue(x, m)
    return argmax_m {val[m]}
```

The algorithm 1 was further adapted according to the best parametrisation of the hybrid player for the SUECA domain, proposed and detailed in Correia et al. (2017).

5 Creating two characters for two robotic game players

Regarding the goals of the work, we aim at creating two different characters, Emys and Glin, to play the SUECA game. Emys was given a *performance-driven goal orientation*, and as such, its behaviours and social actions are more aligned towards winning the game. Glin, by contrast, was given a *learning-driven goal orientation*; consequently, although Glin strives for its team to win the game, it also focuses on fostering team spirit and providing a good game experience.

The challenges associated with defining the two robotic characters were (1) how to reflect different goal orientations through the social interactions of two distinct robots and (2) how to guarantee, in the case of a group of two humans and two robots, that both robots are aware of and synchronised with the others, respect turn taking, and act naturally in a group of four.

To address the first challenge, both robotic game players use the same agent, described in Sect. 4. However, their utterances distinguish them as two different characters. In other words, their repertoire of dialogues was used to author the characters of Emys and Glin. Therefore, each robotic player has a unique set of utterances (420 per robot) for all the perceived game events during a game session. The total amount is balanced to ensure that neither would be more repetitive than the other. Moreover, they produce behaviours with similar frequencies to ensure that neither would exceed the other in its interaction rate.

Table 2 exemplifies the differences between Emys' and Glin's interactions for the same perceived game states. For Emys, the utterances were built based on a competitive perspective, always in pursuit of the best score. For example, the emotion of joy is triggered when the situation reveals that its team is winning. At the same time, Emys will react with an angry emotion when losing and will consequently blame the others, either the partner or the opponents, for the game result. By contrast, Glin was built with different parameters, leading to a more relational perspective, verbalising more support towards its partner. When its team loses, Glin will respond with a sad emotion, encouraging its partner and fostering hope. Note that Glin also plays competitively, desiring its team to win but assuming more of a supportive role.

Another consideration, due to the fact these characters interact verbally, was providing the robots with different voices. It is crucial for the voices to be easily distinguishable, especially because they are embodied in identical robots. Therefore, we used different male Portuguese voices from the same (TTS) engine to ensure that the two robots had similar voice characteristics in terms of lifelikeness, expressiveness, and quality.

Table 2 Examples of utterances from Emys and Glin

Game state	Emys	Glin
Deal cards	"I only accept aces and sevens in my hand!"	"I hope there are good cards for everyone!"
Self playing	"Watch and learn how this is played."	"I am so proud to be on your team!"
Partner played ([a]other team—my team—high)	"Indeed, these points suit our team better."	"Our team is in sync!"
Opponent's turn	"Play..or we will fall asleep."	"It's you, go ahead!"
Partner's turn	"Don't disappoint me."	"Play with confidence!"
Game end—loss	"This cannot continue like this! You have to play better!"	"No worries, next time we will do better!"
Game end—draw	"With this score, I do not like to play."	"It's a draw... no worries, it's okay."

[a] Appraisal according to Table 1

Finally, we would like to emphasise that both characters played the game using the same search algorithm, parameters and heuristics, which is an important design consideration, as we wanted them both to play equally well when placed in the same situation.

5.1 Interaction in a group

To produce natural interactions among the group of four (two humans and two robots) and considering the fact that both human and robotic players play certain roles (partner and opponent) in the game play scenario, the robotic players must be able to interact with each other in a manner as similar as possible to that in which they interact with human players.

Given that these autonomous robots do not have the capability to understand natural language, other mechanisms had to be implemented to achieve natural, believable, and human-like interactions. One fundamental capability required in this scenario is turn taking. For instance, humans use various sensory stimuli to perceive whether another person is going to speak, immediately establishing an order for the speakers according to each situation. Sometimes, a person will even step down from his or her intention to speak because someone else has already said the same thing or because there is no reason to speak anymore. To mimic this natural synchronisation process, we defined a two-phase handshaking protocol as an explicit communication interface. This protocol includes four messages: (1) to inform of an intention to speak, (2) to respond to an intention to speak, (3) to inform that an utterance has begun, and (4) to inform that an utterance has finished. Each robot can perform an utterance only when it receives a positive response. If it receives a negative response, it must wait and retry message (2) until it receives a positive response. A conflict may arise when a robot receives an intention to speak immediately after having sent the same message, as both robots will then receive a negative response and will both enter a retry loop. To avoid a communication deadlock, the two robots will retry their requests after different periods of time, which are randomly generated with values between 0 and 2 s. The next time, one of them may receive a positive response, and if not, they will continue retrying until a request receives a positive response or until a timeout period of 3 s has expired.

This simple mechanism enables a natural and fluid turn taking mechanism between the two robots. A similar mechanism with the human players would also improve the group interaction, but it was currently ignored due to its complexity. Nevertheless, we carefully avoided having explicit questions in the chosen dialogues. When necessary, we replaced them with rhetorical questions instead, as such utterances provide a rich feeling of interaction without requiring explicit answers. For instance, a robot may say "Did we really lose this game?" (Emys) or "What am I going to play next?" (Glin). This detail can be interpreted as a simplified form of two-way communication that allows humans to engage in a conversation or simply to answer the robots.

6 Study 1: character validation

The first study was conducted to validate the differences between the two created characters, i.e., the more performance-oriented character, Emys, and the more relationship-oriented character, Glin. We expected that Emys would be perceived as more competitive, less helpful and less motivating and as providing less emotional security than Glin.

6.1 Sample

We recruited a total of 30 university students (17 males and 13 females) with ages ranging from 19 to 42 years old ($M = 23.03$; $SD = 4.21$). Among the participants, 56.7% had a high level of expertise in the game, 40% had a moderate level of expertise, and only 3.3% had never played the game before. Regarding previous interactions with the (EMYS), 24 participants had previously interacted with it, and 6 were interacting with it for the first time.

Fig. 1 Experimental setting for Study 1

Each participant was randomly allocated to a session in which three human participants played either with Emys or with Glin. This session lasted approximately 1 h, and the instruments used were an (EMYS) robotic head (Kedzierski et al. 2013), two video cameras to record the interactions, a multi-touch table, and a deck of physical cards with printed fiducial markers that could be recognised by the table.

6.2 Procedure

The participants arrived at the room in groups of three. A researcher received them, explained the rules of the game, and conducted a test game to address any doubts that might arise regarding the game rules. After the explanation, the participants joined either Emys or Glin (chosen randomly) at the table and played a set of 3 games. When finished, the participants were administered a set of questionnaires, filled out the consent form and received a thank-you gift (a movie ticket). We presented the consent form at the end of the experiment so that the participants' interactions during the game would be as natural as possible. If any participant had not given consent, his or her data would have been erased. However, all participants signed the consent form.

6.3 Measures

To represent our sample, demographic information was requested in the questionnaires (gender, age, previous interaction with the robot and level of expertise in the game). In addition, all participants, independently of being the partner or an opponent of the robot, responded to the following questionnaires regarding the robot (Emys /Glin):

- *Competitiveness index* (Smither and Houston 1992), used to measure the level of competitiveness perceived in the robot. This measure is usually treated as being of a dichotomous true/false answer type; however, as our goal was to determine a range from the participants' answers, we measured it on a Likert scale ranging from "totally disagree" to "totally agree". An example of a statement would be "I consider Emys a competitive individual" or "When Emys plays, he likes to keep an eye on the score".
- *McGill friendship questionnaire* (Mendelson and Aboud 1999), using three of its dimensions, namely, help (e.g., "Emys helps me when I need it."), motivation (e.g., "Emys praises me when I do something right.") and emotional security (e.g., "If I was worried, Emys would make me feel better"), with scales ranging from "totally disagree" to "totally agree".
- *Relationship assessment scale* (Hendrick 1988), adapted to the context and used to ascertain the level of quality of the relationship with the robot, ranging from "few" to "a lot" (e.g., "How good was Emys relationship with the players?").
- *Godspeed questionnaire* (Bartneck et al. 2009), using the two dimensions of perceived intelligence and likeability to assess the level of intelligence thought to be given to the robot and its perceived likeability, measured as a semantic differential.

All dimensions were measured on a 6-point Likert scale, and when necessary, items were shuffled to mask their dimensions.

6.4 Results

To understand whether the two characters were perceived differently, statistical analyses were performed. When a normal distribution was present, we performed the Student's t-test for independent samples, and when the normality assumption was not met, we used the Mann-Whitney U test. The means and standard deviations are presented in Table 3.

For the *Competitiveness Index*, Emys was rated higher than Glin, with a statistically significant difference ($t(25) = -4.893$, $\rho < .001$). Notably, Glin also presented a certain level of competitiveness, which was expected since it also had the goal of winning the game. Regarding the *McGill Friendship Questionnaire*, there were statistically significant differences in the three measured dimensions of help ($t(28) = 2.312$, $\rho = .028$), motivation ($t(28) = 3.686$, $\rho = .001$), and emotional security ($t(28) = 3.218$, $\rho = .003$), with Glin presenting higher scores than Emys. On the *Relationship Assessment Scale*, Glin was rated higher than Emys, with a statistically significant difference ($t(28) = 5.514$, $\rho < .001$).

These results confirm that the behavioural manipulation of the goal orientations of both robots was perceived as intended: Emys was seen as more competitive, and Glin was seen as more relationship-driven, with a greater capacity to be helpful and motivating and the ability to provide more emotional security. Moreover, the relationship quality scores were also higher for Glin than for Emys. We additionally evalu-

Table 3 Study 1 results: means and ranks with standard deviations for the questionnaire dimensions comparing the evaluations of the Emys and Glin characters

Questionnaire dimensions	Emys	Glin
Competitiveness index*	4.57 ± 0.40	3.86 ± 0.33
McGill		
Help*	3.78 ± 0.89	4.51 ± 0.81
Motivation*	3.79 ± 1.00	4.95 ± 0.69
Emo. security*	3.26 ± 1.09	4.37 ± 0.77
Relationship quality*	4.41 ± 0.52	5.32 ± 0.38
Godspeed		
Perc. intellig.	4.59 ± 0.74	4.93 ± 0.49
Likeability*	10.70 ± 0.88	20.30 ± 0.88

$*p \leq 0.05$

ated whether the roles of the participants (partner/opponent) had any influence on the scores given to the robots, and we found no statistical significance for all measures, suggesting that the role did not affect the evaluations.

Finally, concerning the findings of the *Godspeed Questionnaire*, there was no significant difference between the two robots in the perceived intelligence dimension ($t(28) = 1.511$, $\rho = .142$). This was somewhat expected since we equipped both robots with the same algorithm for solving the card game. Although the game includes an element of chance and each new game presents different winning probabilities for each team, we can conclude that the intelligence levels of both robots were similarly perceived. However, in the likeability dimension, we found a significant difference, with Glin receiving higher scores than Emys ($U = 40.50$, $\rho = .002$).

In general, it seems that our implementations were perceived by the participants as intended, and Glin was rated as more likeable than Emys. We could now move on to the implementation of both characters at the same time, using the two robots to test which would be the preferred partner.

7 Study 2: choosing a robotic partner

The purpose of this study was to assess the participants' preferences regarding the choice of a robotic partner.

7.1 Sample

For the second study, we recruited a new sample consisting of a total of 61 participants (59 university students and 2 workers), 38 male and 23 female, with ages ranging from 17 to 32 years old ($M = 23.66$, $SD = 3.24$). The majority of the participants had never interacted with a robot before and had a moderate or high level of expertise in the game.

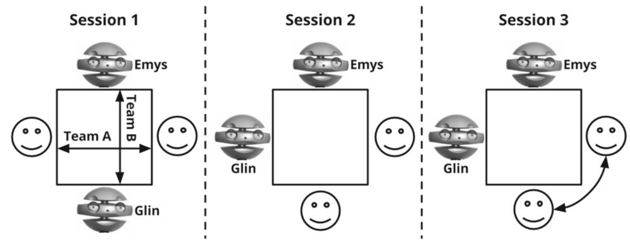

Fig. 2 Experimental setting for Study 2

We measured the level of competitiveness of each participant using the Competitiveness Index (Smither and Houston 1992): 15 participants presented low levels of competitiveness (less than or equal to $M = 3.50$), 36 participants presented some level of competitiveness, and 10 participants showed high levels of competitiveness (higher than $M = 4.50$).

Each session was run with two human participants who did not know each other beforehand. We controlled for this factor to ensure that the participants were in the same position with respect to both each other and the robots. Each session lasted approximately 1 h 30 m, and the instruments used were the same as in the previous study except that two (EMYS) robotic heads were used simultaneously during the game interaction (a name tag was placed below each robot with its name, i.e., Emys or Glin, to allow the participants to easily identify them).

7.2 Procedure

The participants arrived at the room and responded to the first part of the questionnaire (see the Measures subsection below); then, a researcher explained the game rules and conducted a test game to address any doubts that might arise. This study was divided into 3 consecutive sessions, as shown in Fig. 2.

1st Session The two participants partnered with each other and played a set of 3 games against the two robots (Emys and Glin), which acted as their opponents in the game. This session served to expose the participants to the two different characters while having the same role towards each one. After completion, the participants responded to the second part of the questionnaire.

2nd Session Each participant partnered with one of the robots (see Fig. 3), and the group played another set of 3 games. The participants then responded to the third part of the questionnaire.

3rd Session The participants played their last set of 3 games, now partnering with the robots with which they had not played before, and then responded to the fourth part of the questionnaire. At the end, they were given the consent form and were thanked for their participation with a movie ticket.

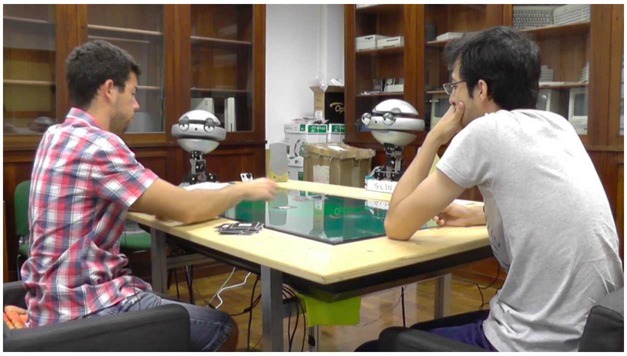

Fig. 3 Experimental setting for Study 2 when each robot was partnering with a human

The balance between the orderings was ensured by the fact that participants attended in pairs and that while one participant was Glin's partner, the other one was Emys' partner. In the last two sessions of each experiment, the participants had the opportunity to partner with both robots. We randomised which participant partnered with each robot first.

7.3 Measures

We used the same questionnaires as in the first study, organised in the following way:

First Part The participants filled out some demographic questions and then an assessment of the *Competitiveness Index* related to themselves.

Second Part The participants completed a questionnaire assessing the two *Godspeed* dimensions (perceived intelligence and likeability) for both robots and answered the following question: "If you could choose one of the robots as your partner, which one would it be? (Emys or Glin)".

Third Part Each participant completed a questionnaire assessing the two *Godspeed* dimensions, the three *McGill Friendship* dimensions (help, motivation and emotional security) and the *Relationship Assessment Scale* with respect to the robot he or she had just partnered with.

Fourth Part The same as the third part of the questionnaire but with respect to the new robotic partner. At the end, the participants were again asked to choose which robot they would prefer to be partnered with for future games and to justify their choice.

All dimensions were measured on a 6-point Likert scale, and when necessary, items were shuffled to mask their dimensions.

7.4 Results

Below, we present the results of this user study, beginning with how participants perceived each robot. Although we previously checked our manipulation, we repeated the analysis to check if the robots were perceived differently in the new 2-robot and 2-human players setting.

Then, we present the results of the participants' initial choice for the preferred robotic partner. We analysed the effect of participants' competitiveness index on this choice.

Finally, we present the participants' last choice after interacting with both robots. We analysed the effect of other measures, such as participants' competitiveness index and the team performance. We also explored changes from the initial to the last choice and the participants' justifications for the chosen partner.

Due to the high number of statistical tests performed, we performed a Holm's sequential Bonferroni correction (Holm 1979) to ensure that there were no false positives in our results, and this assumption was met for all the statistical tests.

7.4.1 Results (I): Perception of the robots

We began by analysing how the participants perceived each robot in their initial interactions. When the normality assumption was not met with the Shapiro-Wilk test, we used the Wilcoxon signed-rank test. The means and standard deviations are presented in Table 4.

Regarding the *McGill Friendship Questionnaire*, there were statistically significant differences in the help ($Z = -5.223$, $\rho < .001$), motivation ($Z = -6.066$, $\rho < .001$) and emotional security ($Z = -5.837$, $\rho < .001$) dimensions, with Glin being rated higher than Emys. For the *Relationship Assessment Scale*, there was also a statistically significant difference ($Z = -4.392$, $\rho < .001$), with Glin being rated higher than Emys, representing a higher relationship quality.

These latter two results confirm the successful behavioural manipulation of the robots. After interacting with both robots, the participants seemed to perceive Glin as having a greater capacity for being helpful and motivating and for providing more emotional security compared with Emys. Moreover, the participants perceived Glin as displaying a better relationship quality than Emys. Overall, these results seem to support the more relationship-driven characteristic with which we attempted to endow Glin, demonstrating the successful development and implementation of the two autonomous robots.

The participants assessed the two dimensions of the *Godspeed Questionnaire* for each robot twice, the first time before partnering with either of the robots and having only observed them as opponents and the second time immediately after having partnered with that robot. For the perceived intelligence dimension, we found no statistically significant difference between Glin and Emys in either the first measurement instance ($Z = -.733$, $\rho = .464$) or the second ($Z = -1.491$, $\rho = .136$). Thus, by using the same decision-making algorithm for both robots in this hidden-information card game, we achieved similar levels of perceived intelli-

Table 4 Study 2 results: means and ranks with standard deviations for the questionnaire dimensions comparing the robots Emys and Glin

Questionnaire dimensions	Emys	Glin
McGill		
Help*	3.35 ± 1.08	4.42 ± 1.13
Motivation*	3.15 ± 1.09	4.79 ± 0.90
Emo. security*	2.58 ± 1.14	4.29 ± 1.19
Relationship quality*	3.93 ± 0.89	4.80 ± 0.93
Godspeed		
Perc. intellig. (BP)	4.51 ± 0.86	4.53 ± 0.99
Likeability (BP)*	3.70 ± 1.19	4.28 ± 0.94
Perc. intellig. (AP)	4.40 ± 1.04	4.55 ± 1.13
Likeability (AP)*	3.51 ± 1.35	5.25 ± 0.75

BP before partnering, *AP* after partnering
$^*p \leq 0.05$

gence in both, as intended. For the likeability dimension, there was a statistically significant difference, with Glin receiving higher scores than Emys in both the first measurement instance ($Z = -3.451$, $\rho = .001$) and the second ($Z = -6.224$, $\rho < .001$).

7.4.2 Results (II): Initial choice of robotic partner

The participants were asked to choose which robot they would like to have as a partner immediately after the first session (in which they had both robots as opponents and had partnered only with another human participant). This allowed us to assess the first impressions people had of the robots and how these would guide their choice of partner. The results showed that 38 of the participants would prefer to have Glin as a partner, whereas 22 preferred Emys. Running a chi-square goodness of fit test, we found a statistically significant difference between the participants' choices ($\chi^2(1) = 4.267$, $\rho = .039$), with more people preferring Glin (63.3%) compared with Emys (36.7%). In this stage of the experiment, the robots were on the same team, and as such, the performance of one robot could not be contrasted with the performance of the other. To better understand the participants' choices, we also compared the participants' competitiveness scores based on their chosen robots using the Student's t-test for independent samples, and we found that there was no statistically significant difference between the competitiveness scores of participants who chose Glin and those who chose Emys ($t(58) = 1.242$, $\rho = .219$). This suggests that at this stage, competitiveness did not influence the partnering choice. Therefore, the participants' choices seem to have been guided by the different social behaviours exhibited; in this case, the participants were more drawn to the relational robot (Glin), which, according to the Results (I) section, was perceived as more likeable than Emys. Thus, the findings support our hypothesis that people seem to prefer a friendlier and more relationship-oriented robotic partner. However, we also wished to investigate whether these characteristics would continue to drive the participants' preferences after they had interacted with both robots as partners.

7.4.3 Results (III): Final choice of robotic partner

When asked to choose a robotic partner in the last questionnaire session (after having partnered with both robots), 35 of the participants preferred Glin and 25 preferred Emys (one participant refrained from choosing). Running a chi-square goodness of fit test, we found no statistically significant difference between the participants' choices ($\chi^2(1) = 1.667$, $\rho = .197$). We then investigated the factors driving the participants' choices at this stage of the interaction.

Looking at the levels of competitiveness of the participants and comparing them according to their final choices, we found a statistically significant difference ($t(58) = 2.953$, $\rho = .005$), indicating that the participants who chose Emys also tended to have higher competitiveness scores ($M = 4.21$, $SD = 0.67$) compared with the scores of the participants who chose Glin ($M = 3.73$, $SD = 0.58$). This implies that a participant's own characteristics (being more or less competitive) played a role in his or her choice of robotic partner after interacting with each robot on his or her team over repeated interactions.

Since the participants partnered with both robots, we also considered the possibility that the performance of the team formed with each robot (winning or losing) also affected the partner choice. To investigate this, we calculated the performance of each human–robot team using the summed results of the sessions, i.e., the sum of the points that Glin's team earned in Session 2 + Session 3, independently of its human partners, compared with the points earned by Emys' team. We observed that based on this criterion, Emys' team won 16 times and Glin's team won 12 times (4 draws occurred). Although this difference was not statistically significant ($\chi^2(1) = .571$, $\rho = .450$), we found a significant association with the partnering preference using Fisher's exact test ($\rho = .008$). It seems that the participants aligned their choices with the robot that was winning more. However, we must be careful with this assumption; each robot was always playing on a team, so if a particular robot won, its win was due not only to its own performance but also to its human partner's performance. Therefore, we can speak of the team performance as a factor influencing the partner choice.

Looking only at the participants who changed their choices of robotic partner between the first session and the last, we found a statistical association between the last chosen robot and that robot's team performance according to Fisher's exact test ($\rho = .002$). By contrast, for the participants whose choices did not change, no significant association was found according to Fisher's exact test ($\rho = .409$). This suggests that the participants who changed their choices did so because of the robot's team performance, thereby solidifying the conclusion that the team performance was indeed one factor accounting for the partner choice, but not the only one.

To clarify whether the robot's character had any influence on the participants' choices at this stage, we analysed their justifications for preferring their chosen robots. For this purpose, two coders (who were completely unaware of the purpose of the study) coded the participants' phrases according to the following coding scheme: they coded a response as *relational* if the justification for the choice of robot was more closely related to team spirit or the robot showing a warmer, more motivating, or more supportive attitude toward its partner, and they coded a response as *competitive* if the justification was based on the robot being the best robot, earning more points, or being more competitive either on its own or towards its opponents. This coding scheme was based on the development objectives for the two different characters. The Cohen's kappa value was k = .73 ($\rho < .001$), revealing good agreement between the coders. We found from the analysis that Glin was chosen 26 times with relational justifications and only 9 times with competitive justifications. By contrast, Emys was chosen 21 times with competitive justifications and 4 times with relational justifications. These results suggest that the robots' characters were also perceived by the participants and used to justify their choice, although this was not the only factor considered.

Overall, these results suggest that *team performance*, a *person's level of competitiveness*, and the *robot's character* play a role in a person's choice of a robotic partner after having previously partnered with it.

8 General conclusions

In this work, we explored preferences regarding robotic partners in mixed teams of humans and robots. Moreover, we studied the factors driving the human participants' partnering choices. For this purpose, we developed two autonomous social robots with different characters, i.e., Emys and Glin, a more competitive robot and a more relational robot, respectively. These two autonomous robots interacted in a group with two humans while playing a competitive game. We began by validating that the two robotic characters were, in fact, differently perceived by the participants. Then, we investigated which of them would be chosen by the participants as a partner for future games. The participants were asked which robotic character (Emys or Glin) they preferred at the two following points in time: (1) before having partnered with either robot and (2) after having played with both robots as partners.

The partner choices seemed to be guided by different factors depending on the context of the participants. In the first session, when the participants had experienced both robots as opponents and had not yet created a partner relationship with either, they seemed to choose their partners based solely on character (either the relationship-driven or competitive robot). At that time, Glin, the relational robot, was the preferred partner. This finding confirms our hypothesis, consistent with the study of Porter (2005), that teams whose members prioritise relational features are perceived more positively (e.g., reporting higher levels of supportive behaviour and higher-quality interactions).

However, at the end of the final session, when they had experienced a partner relationship with each robot, the participants' choices became less clear, calling attention to other factors that came into play. It seems that *personal characteristics* and *team performance* took higher precedence when participants had experienced partner-partner relationships with the robots. The participants seemed to be affected by their *own characteristics* in their partner choices, as we observed that participants with higher levels of competitiveness tended to choose the more competitive robot (Emys), whereas the less competitive participants tended to choose Glin. At the same time, although both autonomous robots played the game using the same algorithm and the difference between the numbers of victories achieved by Emys' and Glin's teams was not significant, there was an association between the team performance and the chosen robot. It was observed that the change in participants' choices between the first and last sessions showed a significant association with team performance. Reinforcing this observation, the performance of the team was also a factor in the final choice of the preferred partner. The same association was not observed for the participants who maintained their choices. In addition, the robot's character also seemed to have influenced the choice, as the participants' justifications of their choices were related to the robots' characters. For example, Glin was chosen because it was much more relational, whereas Emys was chosen because it was more competitive.

The second user study, in particular the first session where both participants were opponents to both robotic characters, was carefully designed to expose the characters to the users on an equal footing. We note, however, that the subsequent user choices and preferences might have been different without this initial session. Moreover, our results do not explore ordering effects, which might be interesting to explore in the future.

Nevertheless, these results have important implications for the creation of robotic teammates who can adapt to their human partners' specific characteristics. Consistent with recent findings (Fraune et al. 2017) showing that people perceive multiple robots that act and look the same as more threatening than a diverse group of robots, people's preferences also need to be considered in the creation of mixed human–robot teams. Indeed, as we move towards scenarios featuring interactions among multiple robots and multiple users, the “diversity” of the robots should not only be investigated but also engineered.

9 Human–robot teams

As demonstrated in the literature on (HRI), in the future, much more complex interactions between humans and robots will exist. These interactions will need to be considered and planned in regard to the design of the robots as well as their social capabilities. Robots that are going to collaborate with humans need to be designed accordingly. Accommodating the partner in the interaction can range from security measures in terms of the material used for its body to adapting its functions to complement the human actions.

On the other hand, when we think about contexts in which social capabilities need to be embedded in the robot, other factors seem to be important in its development. If we want to implement a game partner or an opponent (e.g., a robot in an elderly care centre or a school), other factors need to be considered, and performance alone is not enough to bring enjoyment. The user characteristics will also play a huge part in the development of a game character, e.g., should the robot adapt to each person’s level of competitiveness? Additionally, when playing the role of an opponent, which characteristics should the robot have? These are interesting topics that need to be further explored to understand how robots can function alongside humans in our society and how they can help and be more enjoyable during that interaction. Moreover, other dimensions of the interaction between teams of humans and robots should also be addressed as, for instance, recent findings explore socio-emotional support and gaze behaviours (Oliveira et al. 2018). The study presented here only explores a part of these human–robot teams, as it helps to unveil people’s preferences for a partner in a competitive game.

By unveiling the tendency of people to match the robot’s characteristics with their own characteristics, we also contribute to a more general understanding of membership preferences. Such preferences are linked to the perception of coherent mixed groups of humans and robots and, therefore, to the notion of social groups. These types of groups can naturally emerge in social contexts, even in scenarios without explicit competition (Eyssel and Kuchenbrandt 2012), and are associated with the social identity theory. Previous findings in the (HRI) field have indeed showed that team members with strong levels of group identification also trust more their robotic team partners (Correia et al. 2018). Therefore, the preferences people have when forming teams with robots also mirror their perception of a coherent team and may be related with stronger levels of group identification. Nevertheless, further exploration is necessary to analyse the effect of congruent characteristics, e.g., goal orientation and/or other variables, of the group members on other factors required by the broader field of human–robot collaboration, such as the group performance.

Acknowledgements This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT-UID/CEC/500 21/2013), through the Project AMIGOS (PTDC/EEISII/7174/2014), and through the Project LAW TRAIN (Ref. H2020-FCT-2014/653587). Filipa Correia, Sofia Petisca, and Patrícia Alves-Oliveira acknowledge their FCT Grants (Refs. SFRH/BD/118031/2016, SFRH/BD/118013/2016, and SFRH/BD/110223/2015, respectively).

References

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71–81.

Bornstein, G., & Yaniv, I. (1998). Individual and group behavior in the ultimatum game: Are groups more “rational” players? *Experimental Economics*, *1*(1), 101–108.

Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human–robot teamwork. In: *2005 IEEE/RSJ international conference on intelligent robots and systems* (pp. 708–713). IEEE.

Buro, M., Long, J. R., Furtak, T., & Sturtevant, N. R. (2009). Improving state evaluation, inference, and search in trick-based card games. In: *IJCAI* (pp. 1407–1413).

Chang, W. L., White, J. P., Park, J., Holm, A., & Šabanović, S. (2012). The effect of group size on people’s attitudes and cooperative behaviors toward robots in interactive gameplay. In: *RO-MAN, 2012 IEEE* (pp. 845–850). IEEE.

Coradeschi, S., & Saffiotti, A. (2006). Symbiotic robotic systems: Humans, robots, and smart environments. *IEEE Intelligent Systems*, *21*(3), 82–84.

Correia, F., Alves-Oliveira, P., Maia, N., Ribeiro, T., Petisca, S., Melo, F. S., et al. (2016). Just follow the suit! trust in human–robot interactions during card game playing. In: *2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN)* (pp. 507–512). IEEE.

Correia, F., Alves-Oliveira, P., Ribeiro, T., Melo, F. S., & Paiva, A. (2017). A social robot as a card game player. In: *13th AAAI conference on artificial intelligence and interactive digital entertainment*.

Correia, F., Mascarenhas, S., Prada, R., Melo, F. S., & Paiva, A. (2018). Group-based emotions in teams of humans and robots. In: *Proceedings of the 2018 ACM/IEEE international conference on human–robot interaction* (pp. 261–269). ACM.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040.

Eison, J. A. (1979). The development and validation of a scale to assess differing student orientations towards grades and learning. PhD thesis, University of Tennessee, Knoxville.

Eyssel, F., & Kuchenbrandt, D. (2012). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology*, *51*(4), 724–731.

Fraune, M. R., Kawakami, S., Šabanović, S., De Silva, P. R. S., & Okada, M. (2015). Three’s company, or a crowd?: The effects of robot number and behavior on HRI in Japan and the USA. In: *Proceedings of robotics: Science and systems*. Rome, Italy. https://doi.org/10.15607/RSS.2015.XI.033.

Fraune, M., Nishiwaki, Y., Šabanović, S., Smith, E., & Okada, M. (2017). threatening flocks and mindful snowflakes: How group entitativity affects perceptions of robots. In: *International conference on human–robot interaction, HRI*. ACM Press (to appear).
Gates, B. (2007). A robot in every home. *Scientific American*, *296*(1), 58–65.
Ginsberg, M. L. (2001). Gib: Imperfect information in a computationally challenging game. *Journal of Artificial Intelligence Research*, *14*, 303–358.
Groom, V., & Nass, C. (2007). Can robots be teammates? Benchmarks in human–robot teams. *Interaction Studies*, *8*(3), 483–500.
Hendrick, S. S. (1988). A generic measure of relationship satisfaction. *Journal of Marriage and the Family*, *50*(1), 93–98.
Hinds, P. J., Carley, K. M., Krackhardt, D., & Wholey, D. (2000). Choosing work group members: Balancing similarity, competence, and familiarity. *Organizational Behavior and Human Decision Processes*, *81*(2), 226–251.
Hoffman, G., & Breazeal, C. (2007). Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In: *Proceedings of the ACM/IEEE international conference on Human–robot interaction* (pp. 1–8). ACM.
Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.
Jung, M. F., Lee, J. J., DePalma, N., Adalgeirsson, S. O., Hinds, P. J., & Breazeal, C. (2013). Engaging robots: easing complex human–robot teamwork using backchanneling. In: *Proceedings of the 2013 conference on computer supported cooperative work* (pp. 1555–1566). ACM.
Kedzierski, J., Muszyński, R., Zoll, C., Oleksy, A., & Frontkiewicz, M. (2013). Emys—Emotive head of a social robot. *International Journal of Social Robotics*, *5*(2), 237–249.
Kuchenbrandt, D., Eyssel, F., Bobinger, S., & Neufeld, M. (2013). When a robot's group membership matters. *International Journal of Social Robotics*, *5*(3), 409–417.
Lee, K. M., Peng, W., Jin, S. A., & Yan, C. (2006). Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of Communication*, *56*(4), 754–772.
Mendelson, M. J., & Aboud, F. E. (1999). Measuring friendship quality in late adolescents and young adults: Mcgill friendship questionnaires. *Canadian Journal of Behavioural Science*, *31*(2), 130.
Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103.
Oliveira, R., Arriaga, P., Alves-Oliveira, P., Correia, F., Petisca, S., & Paiva, A. (2018). Friends or foes? Socioemotional support and gaze behaviors in mixed groups of humans and robots. In: *Proceedings of the 2018 ACM/IEEE international conference on human–robot interaction* (pp. 279–288). ACM.
Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.
Porter, C. O. (2005). Goal orientation: Effects on backing up behavior, performance, efficacy, and commitment in teams. *Journal of Applied Psychology*, *90*(4), 811.
Reeves, B., & Nass, C. (1996). *How people treat computers, television, and new media like real people and places*. Cambridge: CSLI Publications and Cambridge.
Rus, D., Donald, B., & Jennings, J. (1995). Moving furniture with teams of autonomous robots. In: *Proceedings of 1995 IEEE/RSJ international conference on intelligent robots and systems 95. 'Human robot interaction and cooperative robots'* (Vol. 1, pp. 235–242). IEEE.
Russell, S., Norvig, P., & Intelligence, A. (1995). *A modern approach. Artificial intelligence* (Vol. 25, p. 27). Egnlewood Cliffs: Prentice-Hall.
Shah, J., Wiken, J., Williams, B., & Breazeal, C. (2011). Improved human–robot team performance using Chaski, a human-inspired plan execution system. In: *Proceedings of the 6th international conference on human–robot interaction* (pp. 29–36). ACM
Smither, R. D., & Houston, J. M. (1992). The nature of competitiveness: The development and validation of the competitiveness index. *Educational and Psychological Measurement*, *52*(2), 407–418.
Sturtevant, N. R. (2008). An analysis of UCT in multi-player games. In: *International conference on computers and games* (pp. 37–49). Springer.
Tapus, A., Ţăpuş, C., & Matarić, M. J. (2008). User–robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, *1*(2), 169–183.
Walters, M. L., Dautenhahn, K., Te Boekhorst, R., Koay, K. L., Kaouri, C., Woods, S., et al. (2005). The influence of subjects' personality traits on personal spatial zones in a human–robot interaction experiment. In: *IEEE international workshop on robot and human interactive communication, 2005. ROMAN 2005* (pp. 347–352). IEEE.

Filipa Correia received a M.Sc. in Computer Science from University of Lisbon, Portugal, 2015. She is currently pursuing a Ph.D. in Human–Robot Interaction and Artificial Intelligence applied to games at University of Lisbon, Portugal, and is a teaching assistant in the courses of Artificial Intelligence, Multi-Agents Systems, and Social Robots & Human–Robot Interaction. She had also been part of the EU FP7 EMOTE project.

Sofia Petisca took a Master degree in Psychology of Emotions in 2013 at Instituto Universitário de Lisboa (ISCTE-IUL), since 2014 she has been working in the field of Human–Robot interactions at GAIPS and INESC-ID, having been part of the EU FP7 EMOTE project. Recently, she has started her Ph.D. with the topic of Morality in human- robot interactions, with the aim to understand how can we better equip our social robots to promote more honest behaviors from people.

Patrícia Alves-Oliveira is a Ph.D. student at ISCTE-IUL in the field of psychology applied to HRI. Pior to her Ph.D., Patrícia has been involved in the study of empathy in robots, and how the empathic capabilities embedded in robotic agents can foster and influence learning amongst children. Patrícia has worked as a researcher in the EU FP7 EMOTE project, developing an empathic robotic tutor aimed at supporting children in the acquisition of curricular topics. In her Ph.D., Patrícia is studying how we can use robots to boost creativity. More specifically, she is interested in how robots can be used to stimulate creative abilities in children and ultimately contribute to a new generation of children that can more easily adapt to creative and innovative societies.

Tiago Ribeiro is an eclectic researcher in pursuit of harmony between arts and interactive technology. He obtained his BSc and MSc in Computer Science and Engineering at Instituto Superior Técnico (IST) with a dissertation about realtime animation of a highly articulated virtual human skeleton with biomechanically-based movement constraints in 2011. After that he has worked as an assistant researcher in the EU FP7 LIREC and EMOTE projects. In both projects his focus shifted from virtual characters, to achieving believable animation and expression in autonomous social robots. As a Ph.D. student at Instituto Superior Técnico, University of Lisbon since 2014, he is focused on discovering how animation artists can take part in the process of developing intelligent animation that autonomous social robots can use while interacting with humans and the external environment, following on principles of traditional animation. His latest work includes designing and building the Adelino craft robot and developing a new expressive kinematics system that adds postural control to an inverse kinematics technique. He has been collaborating with researchers and Ph.D. students in both INESC-ID, Carnegie-Mellon University and Yale University, and has also been active on the organization of academic events such as the HRI Pioneers Workshop 2016 (general chair) and the AAAI Fall Symposium 2016 (on AI for HRI). He has authored and co-authored over 30 peer-reviewed scientific papers, published and distinguished at conferences such as the ACM/IEEE SIGGRAPH (Student Competition Finalist), the ACM/IEEE HRI (Best Paper recipient and nomination), AAMAS, IVA and others.

Francisco S. Melo is a researcher as GAIPS, INESC-ID and an Associate Professor at the Computer Science Department of Instituto Superior Técnico - University of Lisbon (IST), where he teaches courses on Decision making and AI. Since his Ph.D. in 2007, he held appointments in the Computer Vision Lab of the Institute for Systems and Robotics, in Lisbon (2007) and in the Computer Science Department of Carnegie Mellon University (2008–2009). His research addresses problems within decision making and machine learning, particularly on reinforcement learning, planning under uncertainty, multiagent and multi-robot systems, developmental robotics, and human–robot interaction. He has been coordinator of the MAIS-S and INSIDE projects, both involving human–robot interaction in networked robot systems, and participated as a researcher in a number of national and European projects in human robot interaction.

Ana Paiva is a Full Professor in the Department of Computer Engineering at Instituto Superior Técnico (IST) from the University of Lisbon and Coordinator of GAIPS – "Intelligent Agents and Synthetic Characters Group", at INESC-ID. She has founded and leads the GAIPS research group at INESC-ID (http://gaips.inesc-id.pt/gaips/). The group investigates the creation of complex systems using an agent-based approach, with a special focus on social agents. Professor Paiva is recognised worldwide for her contributions in the areas of virtual agents, multi-agent systems, social robotics and affective computing. She has coordinated the participation of INESC-ID and IST in more than 20 European and national research projects in the areas of her expertise, edited six books, and published more than two hundred papers in scientific journals and conferences. Her main research focuses on the problems and techniques for creating agent-based systems (both virtual or robotic) that can simulate human behaviour or establish natural interactions with humans.