Evaluating the Effects of Limited Perception on Interactive Decisions in Mixed Robotic Domains

Aris Valtazanos
School of Informatics, University of Edinburgh
Edinburgh EH8 9AB, United Kingdom
a.valtazanos@sms.ed.ac.uk

Subramanian Ramamoorthy
School of Informatics, University of Edinburgh
Edinburgh EH8 9AB, United Kingdom
s.ramamoorthy@ed.ac.uk

***Abstract*—Many robotic applications feature a mixture of interacting teleoperated and autonomous robots. In several such domains, human operators must make decisions using very limited perceptual information, e.g. by viewing only the noisy camera feed of their robot. There are many interaction scenarios where such restricted visibility impacts teleoperation performance, and where the role of autonomous robots needs to be reinforced. In this paper, we report on an experimental study assessing the effects of limited perception on human decision making, in interactions between autonomous and teleoperated NAO robots, where subjects do not have prior knowledge of how other agents will respond to their decisions. We evaluate the performance of several subjects under varying perceptual constraints in two scenarios; a simple *cooperative* task requiring collaboration with an autonomous robot, and a more demanding *adversarial* task, where an autonomous robot is actively trying to outperform the human. Our results indicate that limited perception has minimal impact on user performance when the task is simple. By contrast, when the other agent becomes more strategic, restricted visibility has an adverse effect on most subjects, with the performance level even falling below that achieved by an autonomous robot with identical restrictions. Our results could inform decisions about the division of control between humans and robots in mixed-initiative systems, and in determining when autonomous robots should intervene to assist operators.**

***Index Terms*—Interactive teleoperation; limited perception.**

I. Introduction

Most existing robotic systems that are deployed in field applications (e.g. rescue robot teams, de-mining, unmanned aerial vehicles) depend on teleoperation. Many such domains of interest are of a ***mixed*** nature, i.e. feature both fully autonomous robots and robots teleoperated by humans (or physically present humans). Interaction and coordination between such heterogeneous agents is a challenging task, largely due to their varied actions, perception, and cognitive capabilities.

When looking at how humans (tele)operate in mixed domains, it is important to assess how these heterogeneous capabilities affect their ability to make robust *decisions*, in the presence of other, possibly adversarial, interacting agents. Humans are generally believed to have a superior grasp of context and situational awareness than autonomous robots. This is one reason why most deployed systems still depend quite heavily on the human user to control robots. However, this awareness also depends on the *perceptual* information made available to operators, which influences how they perceive their own robot's surroundings and the state of other interacting robots.

In many realistic situations, this information may be sparse or incomplete; for example, an operator controlling a rescue robot in a disaster site may only have access to the robot's noisy camera feed. Thus, a subject having *full* visibility of the environment may be able to fully understand how other agents are behaving, and plan the actions of the teleoperated robot accordingly. By contrast, if the same person has *limited* visibility of the environment, the decisions may be less informed and thus less effective. In the latter case, where users are effectively constrained to have the same perceptual capabilities as a robot, it is unclear whether their decisions would be able to exceed, or even match, the performance level of an autonomous agent. This is an important issue to be addressed in systems where the autonomous system can intervene to assist the human partner.

In this paper, we consider the problem of human-robot interaction in perceptually constrained mixed robotic domains, and present empirical data addressing the following questions:

- What is the effect of incompleteness and asymmetry of information on human teleoperation performance in interactive robotic tasks?
- Where should the boundary between human control and autonomy lie, and what is the correlation between the effects of perceptual limitations and the strategic content of interactive tasks?

We view these issues as central not only to understanding the factors that influence interactive decisions, but also to designing mixed robotic systems that can successfully combine the relative merits of human control and autonomy.

In order to address the above questions, we evaluate the performance of several subjects in two different interactive tasks involving a teleoperated and an identical autonomous NAO humanoid robot. Both tasks share the following properties:

- The human subjects do not know a priori how the autonomous robot will behave, nor can they exchange any data with it during the task. Thus, they can only infer its decisions through *observation* and *repeated interaction*.
- The tasks are *fully interactive*, requiring subjects to make several decisions over a short time horizon and also to *respond* to the actions of the autonomous robot.

The first task is a *cooperative* target allocation task, where the two robots must reach two different targets without interfering with each other. The second is an *adversarial* task, where the two robots compete in a soccer penalty shoot-out (autonomous striker vs. teleoperated goalkeeper). This task is considerably harder for two *interrelated* reasons:

- The autonomous robot is a ***strategic adversary*** who seeks to outperform the human through deceptive manoeuvres.
- The human subject must estimate and infer finer-grained information, e.g. the absolute states of the robots and the most likely kicking direction selected by the striker.

In both cases, we first evaluate subjects under full observability of the interaction environment, and we subsequently constrain them to viewing only a live feed from the robot's camera.

Main hypothesis: In light of the above constraints, our core hypothesis is that only a small proportion of subjects should perform worse in the cooperative task under restricted perception, whereas a greater fraction would be impacted in the adversarial task under these conditions. In other words, we hypothesise that the *combined challenge* of reasoning about *absolute states* and the *strategic behaviour* of the adversary will have an adverse effect on human performance under limited visibility in the second task, unlike the simpler interaction and inference requirements posed by the first task.

In the remainder of this paper, we first review related work from the robotics literature (Section II). We then describe the experimental setup and the interactive tasks (Sections III-IV). In Section V, we present empirical results of our evaluation on several subjects. Our results suggest that restricted visibility is more likely to impact subjects in strategic interactions, where there is greater uncertainty over the autonomous robot. We review our key contributions in Section VI.

II. Related Work

Human-robot interaction is often evaluated in the context of cooperative tasks, where the interacting parties must collaborate to achieve a common goal (e.g., cooperative object manipulation [8][10]). Many such interactions are centred around the ability of the robot to receive and follow instructions from a human, in order to fulfil its role in the task (e.g., [13][18]). Furthermore, several collaboration studies are concerned with modeling human intentions; various approaches have been proposed to this effect, such as velocity-based impedance control [9], dynamic Bayesian networks [17], or interaction history records [8]. In our work, we look at the related problem of how humans account for the intent of autonomous robots in dynamic interactions. Moreover, in our experiments, subjects must infer the robot's intent only through observation which is progressively restricted. Thus, our tasks present different challenges than corresponding problems in the existing literature.

The influence of perception in human-robot cooperation has been previously examined in the context of recognising actions and learning skills from observation (e.g. [11], [12]). Learning from demonstration under perceptual constraints was studied in [7], where it was shown that robots can learn more effectively when the perception of human demonstrators is restricted to be similar to their own. An interesting result in that study was that robots were able to learn more quickly from restricted-perception demonstrations, even though their quality was often inferior to full-perception ones. A similar study on human teleoperation was conducted in [14], where user performance was found to be correlated to the availability of perceptual information. In this paper, we are similarly interested in assessing the effects of constrained visibility on human performance. However, our focus is not on learning from demonstrations provided independently by humans, but instead on evaluating human performance in a purely *interactive* environment, where the human and the robot *simultaneously* engage in cooperative and adversarial tasks.

Fatigue and stress have also been considered as influencing factors during teleoperation. In one recent study [15], subjects were evaluated in a remote grasping task, in consecutive trials totalling up to ten minutes. This study found no clear evidence of performance degradation due to fatigue. In our work, where experiments have a similar duration, we similarly do not find any clear evidence for fatigue-induced effects on performance.

A mixed robotic domain that resembles our setup is Segway soccer [5][6], which involves mixed teams of humans and robots mounted on Segways. Our work shares a similar motivation in that humans and robots engage in a common task using similar physical capabilities. However, we also note some important differences. First, the pace of the interaction in our experiments is faster than Segway soccer (as demonstrated in the supporting video in Section V), requiring more frequent decisions. Second, being one-to-one interactions between teleoperated and autonomous robots, our tasks offer a more direct comparison of human-robot decision-making than Segway soccer, where role allocation is less clear. Thus, although our domain features fewer robots, it ensures that humans and robots get equal interaction time; by contrast, in a mixed-team game, humans could potentially supplant the role of the robots. Third, Segway soccer captures only the full visibility case of our tasks, as humans perceive the world through their own eyes. Here, we seek to further constrain the interaction by restricting the perception of the subjects.

III. Experimental Setup

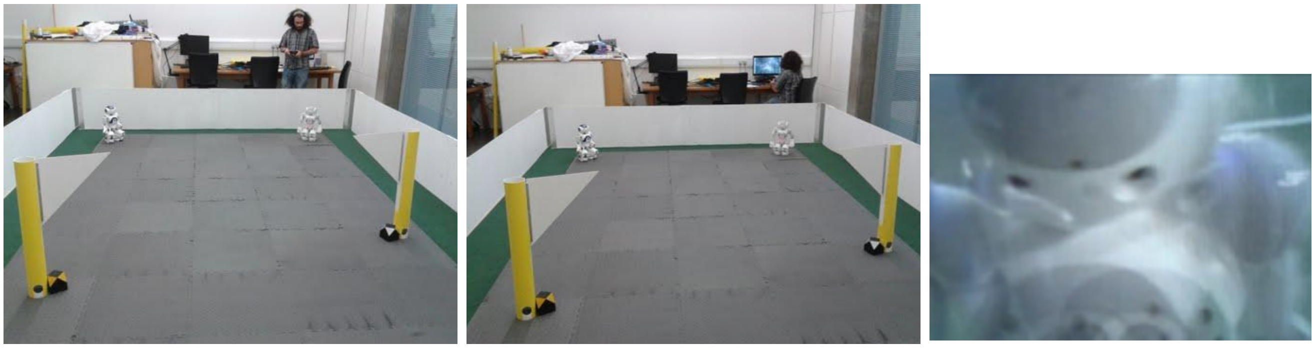
A. Humanoid robot

We use the NAO robot, shown in Figure 1(a), a 58cm-tall humanoid with 21 degrees of freedom [2]. The NAO is the official robot of the RoboCup Standard Platform League (SPL) [1], in which our team, Edinferno [4], is a participating member, having reached the quarterfinals in 2012. Each robot has two cameras capturing images at approximately 30fps. The robot has a limited field of view, but it can move its head in order to track a different area in its environment. Our software is based on the B-Human code release [16], which provides modules for fast walking, vision, and self-localisation.

B. Interaction environment

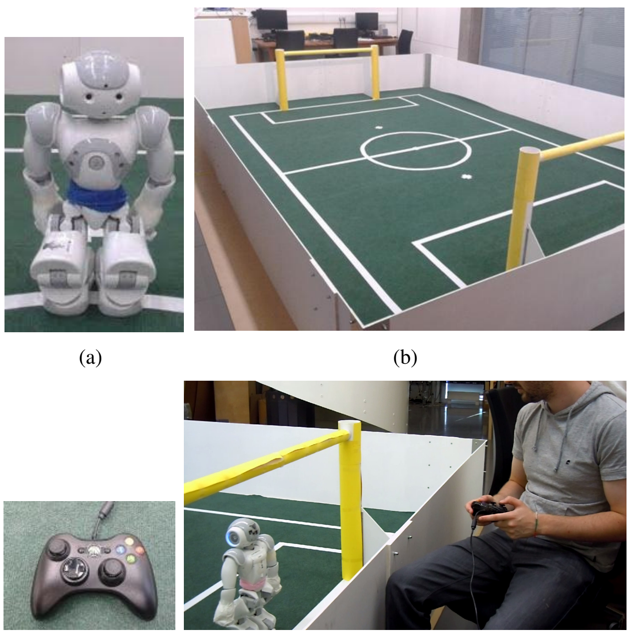
The robots interact in an open arena (Figure 1(b)) modeled on the official SPL field [3], with dimensions of 4.5x3.0m.

In the penalty shooting task, robots use field goals and lines as landmarks, in order to compute their position in the field, and thus determine the relative distance and direction of



(a) (b) (c)

Fig. 2. Experimental setup - cooperative task. (a): An autonomous (top left) and a teleoperated (top right, in front of human operator) robot must reach two different targets (indicated by the flagpoles) without interfering with each other. The autonomous robot randomly selects a target to navigate to, which the subject must infer during the interaction, in order to lead his robot to the other target. The initial positions of the robots are fixed but the locations of the targets change between trials. (b): Same task, but the subject now has access only to the robot's noisy camera feed (shown in (c)).



(a) (b)

(c) (d)

Fig. 1. Experimental setup. (a): The NAO humanoid robot. (b): The arena where the two robots interact. (c): The controller used to command the robot. (d): Remote control of a robot by a subject.

the goal. The goals are 1.40m wide and are painted yellow, whereas lines and field markings are white and placed at specified positions [3]. For the cooperative task, self-localisation is not required, so the field is simply used as an open arena.

C. Teleoperation

Subjects control the robot using an Xbox controller (Figures 1(c)-1(d)). The inputs are converted to action commands (e.g. forward walk) and transmitted wirelessly to the robot.

IV. Interaction Scenarios

A. Cooperative task – Target allocation

In the cooperative target allocation task, the two robots are placed in the arena as shown in Figure 2(a). The task requires the robots to reach two different targets in an arena. The initial positions of the robots are fixed as in Figure 2(a), but the targets are moved around the arena between trials. The autonomous robot initially selects a target at random, and begins moving towards it. The human subject must then infer where the autonomous robot is heading, and steer his own robot to the other target as fast as possible. Subjects must also avoid collisions or interference with the autonomous robot.

The autonomous robot has no external information (e.g. positions from an overhead camera) and perceives the world only through its own perspective camera. There is also no communication between the robots, so there is no prior (or interactive) agreement on the allocation of the targets.

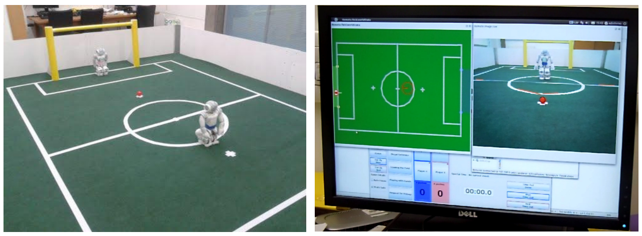
1) Autonomous robot behaviour: The autonomous robot navigates to its chosen target using a simple visual servoing routine. The targets are colour-coded so that they can be easily identified. The random selection of a target is enforced by having the robot initially look away from the arena (so that no targets are visible), and then randomly select whether it should start turning left or right. The robot then keeps turning until it locates a target, and then starts moving towards it.

2) Full vs. restricted perception: We consider the cooperative navigation task in two situations. In the first case (Figure 2(a)), the subject may view the entire arena, thus having full visibility of the environment. In the second case (Figure 2(b)), the subject is restricted to viewing only the teleoperated robot's noisy camera feed (Figure 2(c)) on a computer screen. Thus, the subject is constrained to have the same perceptual capabilities as the autonomous robot, so the two robots differ only at the behavioural level (autonomous vs teleoperated).

In the full visibility case, subjects have a clear view of both robots and both targets. Thus, it is relatively straightforward to identify where the autonomous robot is heading, and, assuming adequate familiarity with the joystick controller, lead the teleoperated robot to the appropriate destination. However, when perceptual information is restricted, recognising the autonomous robot's behaviour and steering the teleoperated robot becomes more challenging.

3) Teleoperation commands: The subject may control the translational (forward-backward-side steps) and rotational (turn left-right) motion of the robot. In restricted visibility, there are additional inputs to control the robot's head movement and scan different parts of the world through its camera.

B. Adversarial task – Penalty shooting



(a) (b)

Fig. 3. Experimental setup - adversarial soccer penalty shooting task. (a): Initial poses of the autonomous striker (near side, blue waistband) and the teleoperated goalkeeper (far side, pink waistband). (b): Restricted perceptual information. Left: Visualisation of the robots' self-localisation estimate (shown by the red markings on the field drawing). Right: Perspective view of the goalkeeper, looking at the ball and the approaching striker.

The adversarial task is a penalty game between an autonomous striker and a teleoperated goalkeeper. The initial positions of the robots are shown in Figure 3(a). The game loosely follows the rules of SPL penalty shooting [3]. The striker has one minute to score a goal and is allowed one kick per trial, so ball dribbling is not permitted. The goalkeeper may not leave or touch the ball outside the penalty box; such violations result to a goal awarded to the striker. The striker has access to a single, straight kick; thus, to shoot towards the goal edges, it must adjust its orientation accordingly.

The objective for the human is to guess which way the autonomous striker is going to shoot, and move the goalkeeper to a suitable shot-blocking position. This is considerably harder than cooperative navigation, as the autonomous robot now attempts to *outperform* the human, by strategically trying to score a goal. Thus, the human must also continuously reason about the absolute positions of the robots in the field.

1) Autonomous robot behaviour: In contrast to the cooperative task, where only relative distances to the targets are required, the autonomous robot now determines its absolute position in the field. As no external information is provided, the robot processes the images retrieved from its camera to identify the relative positions of various landmarks, such as goal posts and field lines. This information is passed to a self-localisation module, which computes the robot's absolute pose (position and orientation) through a particle filter.

The behaviour of the striker was programmed from human demonstration examples. We recorded the control inputs of several subjects controlling a teleoperated striker against an autonomous goalkeeper (the dual of the problem we are considering here). The autonomous goalkeeper followed a simple heuristic algorithm, where the blocking position on the goal line is chosen based on the observed orientation of the striker. Demonstrations were labeled as being either successful (a goal was scored) or unsuccessful (miss). The autonomous striker was programmed to use a probabilistic mixture of the successful demonstrations against the teleoperated goalkeeper.

2) Full vs. restricted perception: Under restricted visibility, subjects are now provided with both the robot's live camera feed and a visualisation of the two robots' self-localisation estimates (Figure 3(b)). In the full visibility case, uncertainty in localisation presents the autonomous robot with an even greater perceptual handicap than in the cooperative task, as noisy or incorrect positional information is likely to lead the striker to misinformed decisions on its adversary. For humans, restricted visibility introduces the challenge of inferring the absolute positions of the robots, using only noisy sensory data.

3) Teleoperation commands: As in the previous task, subjects may control the translational and rotational motion of the goalkeeper. There are also inputs for spreading the robot's legs to block the ball. The goalkeeper is programmed to track the ball and the approaching striker automatically (as in Figure 3(b)), removing the need to control the robot's head separately.

V. Results

We evaluated the two tasks on 40 different subjects; 20 of these subjects were tested just on the adversarial task, 10 just on the cooperative task, and 10 participants on both tasks. The experimental sample was varied, consisting of both male and female subjects, young children and adults, users with previous robotics experience and users who were interacting with robots for the first time. Snapshots from recorded trials are given in Figures 4 and 5. Further examples are available in the supporting video of this paper (available online: http://www.youtube.com/watch?v=6xi7WPgg46A).

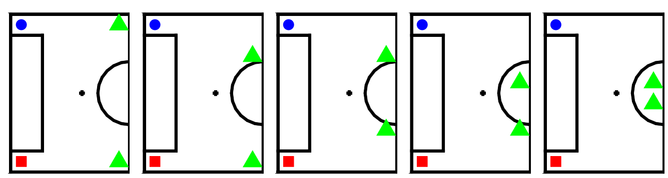


Fig. 6. The five target configurations, cooperative task. *Blue circle*: Teleoperated robot initial position. *Red square*: Autonomous robot initial position. *Green triangles*: Target positions.

For the target allocation task, each subject was evaluated on 5 different target configurations, which are shown in Figure 6. Targets were progressively moved closer to increase the difficulty of the task. Subjects were initially tested on each configuration under full visibility, and then they were asked to repeat this procedure viewing only the robot's camera feed. In each trial, we recorded the targets selected by the robots, the time taken by the teleoperated robot to reach the selected target, whether or not there was a collision with the autonomous robot, and the subject's joystick inputs. As target positions were known in each trial, we divided the distance to the selected target with the total time taken by the subject, to obtain the *average speed* as a normalised performance metric.

For penalty shooting, subjects controlled the goalkeeper for 5 trials under full visibility, and then for a further 5 trials

Fig. 4. Cooperative task. Navigation targets are now represented as orange balls. *Top:* A subject controlling the robot (blue waistband, starting at the right) under full visibility. *Middle:* A trial as seen through the teleoperated robot's camera. *Bottom:* The same subject controlling the robot under restricted visibility.

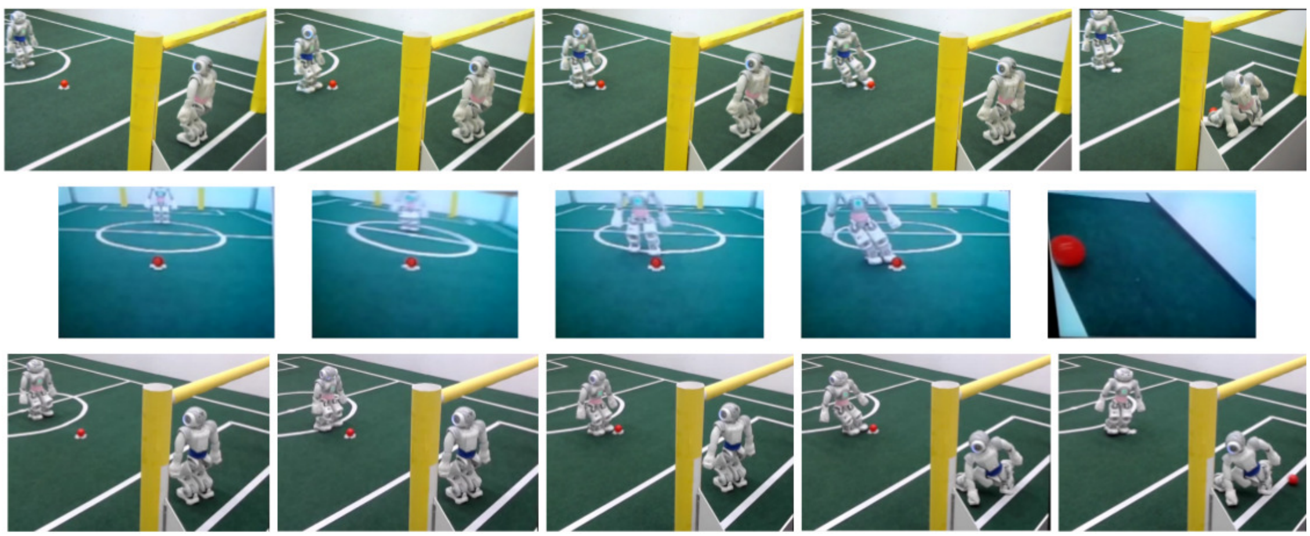


Fig. 5. Adversarial task. *Top:* Full visibility - a teleoperated goalkeeper (operator not shown) saves a shot. *Middle:* A trial as seen through the robot's camera. The last snapshot shows the view of the goalkeeper after an unsuccessful dive to save the ball. *Bottom:* Limited visibility - different subject conceding a goal.

under limited visibility. We recorded the outcome of each trial (goal/no goal), the control inputs of the subject, and the self-localisation estimates of the two robots during the trial.

A. Overall performance

1) Performance metrics: Results for the overall metrics (average speed for target allocation, goals conceded for penalty shooting) are shown in Figure 7. For target allocation, there was little difference between visibility conditions, in both successful execution rate (collisions with autonomous robot) and performance rate (average speed). An interesting pattern is observed in the subject-specific illustration of the results (Figure 7(c) - left), where there is a roughly equal number of subjects with improved and deteriorated performance between the two visibility cases. This suggests that reduced visibility is not an impeding factor in this simple interactive setting.

By contrast, most subjects appeared to struggle more under restricted visibility in the adversarial task. About two thirds of the subjects saved fewer goals when this restriction was applied, while only 4 out of 30 managed to save more (Figure 7(c)-right). For this task, we also recorded the distance of the goalkeeper from the optimal blocking position at the time of the shot (Figure 8(a)). Through this metric, we model how well subjects were able to respond to the moves of the autonomous striker, and lead goalkeepers to a position that maximises the chances of a save. As shown in Figure 8(b), the recorded distance for almost half of the subjects increased considerably under restricted visibility.

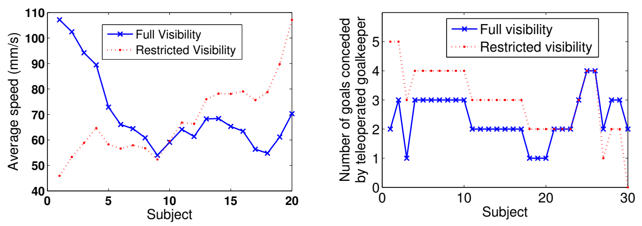
2) Performance rate: Table 7(d) shows a time-indexed representation of the overall results for the different presented experiments. Due to the small number of trials and the short duration of each trial (at most 1 minute in both tasks), subjects

Visibility	Full	Restricted
Mean average speed over all subjects (mm/s)	86.57	76.70
Minimum average speed	38.83	32.31
Maximum average speed	128.22	93.62
Standard deviation of avg. speed	20.05	13.66
Number of collisions with autonomous robot (out of 100 trials)	4	2

(a) Overall performance metrics – cooperative task.

Visibility	Full	Restricted
Total number of goals conceded	71/150	90/150
Mean goals conceded per subject	2.36/5	3/5
Standard deviation	0.81	1.14
Mean goal difference between visib. cases	0.663	
Standard deviation	0.994	

(b) Overall performance metrics – adversarial task.

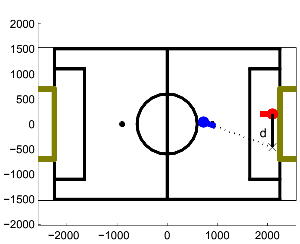


(c) Performance metrics per subject - average speed in target allocation (left), goals conceded in penalty game (right). In each graph, values are sorted by the difference of the performance of the subject between full and restricted visibility. Values towards the left represent subjects most affected by restricted visibility, as indicated by the performance degradation.

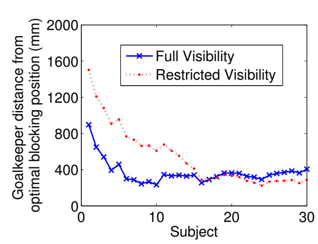
Trial	1	2	3	4	5
Avg. speed (full v.)	69.5	94.3	82.4	89.1	97.4
Avg. speed (restr. v.)	67.6	76.1	77.9	78.1	83.8
Goals conc. (full v.)	0.27	0.67	0.40	0.53	0.50
Goals conc. (restr. v.)	0.60	0.80	0.40	0.50	0.70

(d) Time-indexed representation of overall results – mean values per trial.

Fig. 7. Overall performance metrics – both tasks.



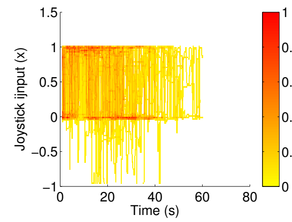
(a)



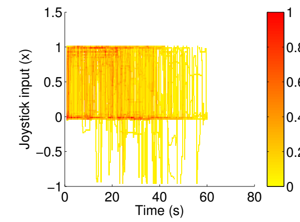
(b)

Fig. 8. Alternative performance metric for adversarial task: distance from optimal blocking position. (a): Explanation of metric. Poses of striker and goalkeeper at time of kick - optimal position for goalkeeper is the intersection of the line formed by the striker's orientation, and the goalkeeper's line of motion. (b): Results per subject, sorted by difference between visibility cases.

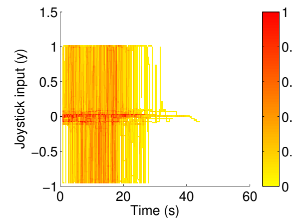
appear not to be affected by factors such as fatigue or stress, which could cause a visible performance degradation in longer experiments. In the restricted visibility instance of target



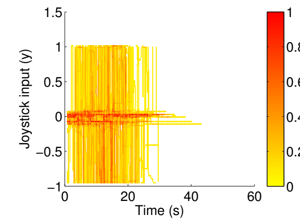
(a) Cooperative - full



(b) Cooperative - restricted

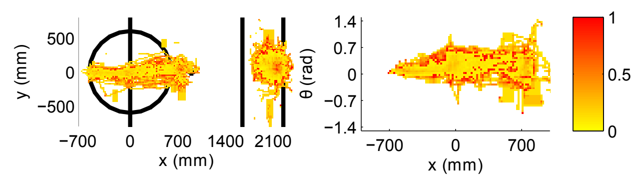


(c) Adversarial - full

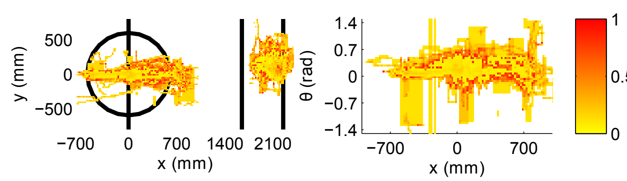


(d) Adversarial - restricted

Fig. 9. Heat maps of recorded user inputs, all trials. Colour indicates the percentage of trials in which a particular control input/time pair was recorded. *Top row:* Cooperative task - forward motion (positive direction: front). *Bottom row:* Adversarial task - side motion (positive direction: left).



(a) Full visibility.



(b) Restricted visibility.

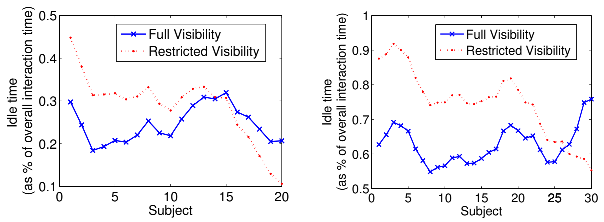
Fig. 10. Heat maps of recorded striker and goalkeeper trajectories, all trials. Colour indicates the percentage of trials in which a particular point was recorded. *Left subplots:* heat maps for forward (x) - side (y) motion trajectory components - left blob corresponds to autonomous striker, right blob to teleoperated goalkeeper trajectories. *Right subplots:* heat maps for forward (x) - rotational (θ) motion components for the striker.

allocation, subjects are seen to improve their performance over time, without however reaching the average speeds attained in the full visibility case. By contrast, there is no conclusive evidence of time-induced learning in the other experiments, with the mean performance fluctuating across different trials.

B. User control inputs and trajectories

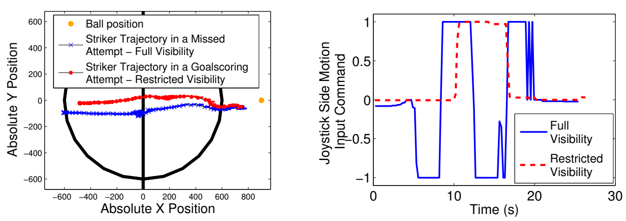
In addition to evaluating overall performance, we compared the variation of user control inputs under the different experimental conditions. Figure 9 provides a heat map representation of all recorded inputs for the two most frequently used axes of motion in the two tasks – the forward motion in target allocation and the goalkeeper's side motion in penalty shooting. In the cooperative task (Figures 9(a)-9(b)), we again observe little variation between full and restricted visibility. However, in the adversarial task (Figures 9(c)-9(d)), the intensity of commanded motion is stronger in the full visibility case.

To further quantify this discrepancy, Figure 10 shows heat maps for all striker and goalkeeper trajectories in the adversarial task. It can be seen that although the trajectories of the autonomous striker are similar in both cases, goalkeepers move towards the edges of the goal less frequently in the second case. This partly explains the higher number of goals conceded by teleoperated robots under restricted visibility.



(a) Cooperative task. (b) Adversarial task.

Fig. 11. Idle times per subject. The idle time is the percentage of the overall time during which no command was sent from the subject to the robot.



(a) Striker trajectories. (b) Goalkeeper control inputs. Positive direction: left.

Fig. 12. Effects of idle time on performance of a specific subject. (a): Two similar trajectories by the striker against this subject, one per visibility case. Only the full visibility attempt was saved by the teleoperated goalkeeper. (b): Illustration of the variation of the subject's side motion between these trials.

Moreover, we looked at how control inputs varied between tasks on a subject-to-subject basis. To this end, we measured the average *idle time*, i.e. the percentage of time during which no command was input by a subject (Figure 11). Idle time is considerably higher in penalty shooting, where subjects spend more time observing the autonomous robot's approach before they move their own robot. However, we also note that both the percentage of subjects whose idle time increases when visibility is restricted, as well as the average rate of this increase, are considerably higher in the adversarial task.

Restricted visibility was also found to impact the *response time* of subjects in the adversarial task. To illustrate this effect, Figure 12(a) two similar autonomous striker trajectories (one for each visibility case) against a subject, and the corresponding user inputs. Although the trajectories are similar, only the full visibility one was saved by the teleoperated goalkeeper. As seen in Figure 12(b), this discrepancy is partly explained by the more delayed response in the restricted visibility case, where the subject needs more time to make sense of the interaction.

C. *Statistical significance*

1) Main hypothesis: In order to assess the statistical significance of our overall results, we tested for the *contradiction* of the main experimental hypothesis as stated in Section I. In other words, our null hypotheses are that a worse performance would be observed for a majority of subjects (greater than 75%) in the simple cooperative task, and for a minority (less than 25%) of subjects in the more complex strategic task. To assess these null hypotheses, we conducted a t-test for the overall performance indicators – the average speed in target allocation, and the number of goals conceded in the penalty game. We measured the percentage of subjects for which performance deteriorated in each case, and computed

$$t = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n}, \tag{1}$$

where $\bar{x}$ is the sampled percentage, μ_0 is the hypothesised percentage (75% in the cooperative task, 25% in the adversarial one), s is the sample standard deviation, and n is the sample size. Based on a two-tailed t-test for the two null hypotheses, we obtain p-values of 0.013 and 0.03, respectively. So, at a 5% significance level, we reject the null hypotheses, and conclude that limited perception does not have a significant impact in the cooperative task, while having a non-negligible effect in the adversarial one which features more severe constraints on perception and action – this was our original main hypothesis.

2) Explaining factors: The difficulty in the adversarial task lies in the determination of the autonomous adversary's strategy, and the estimation of the absolute states of the interacting robots. As these challenges are *not* independent of each other, they cannot be explicitly decoupled in order to assess their individual contribution to the overall difficulty of the task. However, we can assess the correlation between the two secondary metrics, the idle time and the distance from the optimal position, and the overall performance. In defining these metrics, our hypothesis is that increases in each metric between visibility cases are linked to performance degradation, i.e. that the overall mean goal difference (Table 7(b)), and the corresponding difference for subjects impacted by each metric should be comparable (i.e. not differ by more than 1 goal, which is approximately equal to the computed standard deviation).

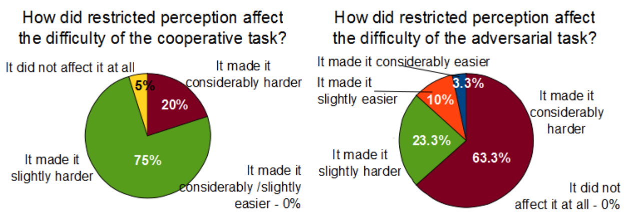
For each metric, we conducted a two-sample pooled t-test to assess its effect on overall performance. We measured statistics for the subjects for which idle time/distance from optimal position increased under limited perception, and computed

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \tag{2}$$

where $\bar{x}_1, s_1, n_1$ are the overall mean goal difference, standard deviation, and sample size, $\bar{x}_2, s_2, n_2$ are the corresponding values for the subset of subjects for which the value of the metric increased, and d_0 is the hypothesised mean difference. We tested for the contradictory null hypotheses that the two metrics cannot be used to explain performance degradation, i.e. that the difference between each $\bar{x}_2$ and the overall mean

$\bar{x}_1$ will be more than 1 goal. For these tests, we obtain p-values of 0.010 for idle times and 0.008 for optimal positions. At a 5% significance level, we reject the null hypotheses, and conclude that an increase in the idle time or the distance from the optimal position is likely to be matched with an increase in the number of conceded goals under restricted visibility.

D. User experiences



(a) Cooperative task. (b) Adversarial task.

Fig. 13. User experiences on restricted visibility.

After each experiment, we asked subjects to give us their opinion on the impact of restricted visibility on their behaviour (Figure 13). In both tasks, most subjects stated that limited perception impacted their performance. However, the dominant response in the first case was that restricted visibility made the task only "slightly harder", whereas most users found the adversarial task "considerably harder". Another interesting result was that some subjects found the penalty game easier under limited perception, with one subject labelling it "considerably easier"; not surprisingly, this is the (only) subject who in Figure 7(c)-right saved all 5 shots under restricted visibility.

VI. Conclusions

Our experimental analysis suggests that limited visibility is more likely to affect teleoperation performance in challenging, adversarial tasks, which require continuous inference of the absolute state and strategy of an interacting robot. Furthermore, restricted perception appears to affect the ability of humans to (inter)act *strategically*, with several subjects being deceived by the autonomous adversarial robot more easily. By contrast, when the task is not particularly challenging and requires only very basic modeling of robot states and strategies, most subjects are likely to be unaffected by this restriction.

Mixed robotic environments are becoming increasingly important in human-robot interaction, as several applications demand an interplay between autonomous and teleoperated agents in complex physical settings. In many such domains (e.g. rescue robotics), perceptual information is inherently limited, so it is important to identify situations where humans might fall short in teleoperating a robot, and how autonomous robots can compensate for these weaknesses. In this respect, our work contributes an empirical evaluation which highlights interaction scenarios and visibility conditions where human control is likely to be problematic, and where autonomous robots can perform more robustly. Our experiment also informs decisions about when to assist human decision makers in teleoperation, and how to structure the balance between human command and robot autonomy. We believe that our methodology can be applied in the design of *mixed robotic teams*, where there is a need to empirically determine both the optimal composition (how many autonomous/how many teleoperated?) of a team, and the roles (what should each autonomous/teleoperated robot do?) of its constituent members.

Acknowledgments

This work has taken place in the Robust Autonomy and Decisions group, School of Informatics, University of Edinburgh. The RAD Group is supported in part by grants from the UK Engineering and Physical Sciences Research Council (EP/H012338/1), the European Commission (TOMSY Grant 270436, FP7-ICT-2009.2.1 Call 6) and a Royal Academy of Engineering *Ingenious* grant. AV has been supported by a doctoral studentship from the Research Consortium in Speckled Computing, funded by the Scottish Funding Council (Strategic Research Development Programme R32329) and EPSRC (Basic Technology Programme C523881).

References

[1] RoboCup Standard Platform League. http://www.tzi.de/spl.

[2] NAO robot datasheet. http://www.aldebaran-robotics.com/en/Discover-NAO/nao-datasheet-h25.htm.

[3] Standard Platform League Rule Book. http://www.tzi.de/spl/pub/Website/Downloads/Rules2012.pdf.

[4] Team Edinferno. http://wcms.inf.ed.ac.uk/ipab/robocup.

[5] B. Argall, Y. Gu, B. Browning, and M. Veloso. The first segway soccer experience: Towards peer-to-peer human-robot teams. In *HRI*, 2006.

[6] B. Browning, L. Xu, and M. Veloso. Skill acquisition and use for a dynamically-balancing soccer robot. In *AAAI*, 2004.

[7] C. Crick, S. Osentoski, G. Jay, and O. C. Jenkins. Human and robot perception in large-scale learning from demonstration. In *HRI*, 2011.

[8] P. F. Dominey, G. Metta, F. Nori, and L. Natale. Anticipation and initiative in human-humanoid interaction. In *IEEE-RAS International Conference on Humanoid Robots*, pages 693–699, 2008.

[9] V. Duchaine and C. M. Gosselin. General model of human-robot cooperation using a novel velocity based variable impedance control. In *EuroHaptics Conference*, 2007.

[10] A. Edsinger and C. Kemp. Human-robot interaction for cooperative manipulation: Handing objects to one another. In *IEEE International Symposium on Robot and Human interactive Communication.*, 2007.

[11] M. Johnson and Y. Demiris. Perceptual perspective taking and Action Recognition. *I. J. of Advanced Robotic Systems*, 2(4):301–308, 2005.

[12] S. Lallée, S. Lemaignan, A. Lenz, C. Melhuish, L. Natale, S. Skachek, T. van der Zant, F. Warneken, and P. F. Dominey. Towards a platform-independent cooperative human-robot interaction system: I. perception. In *IROS*, pages 4444–4451, 2010.

[13] S. Lallée, E. Yoshida, A. Mallet, F. Nori, L. Natale, G. Metta, F. Warneken, and P. F. Dominey. Human-robot cooperation based on interaction learning. In *From Motor Learning to Interaction Learning in Robots*, pages 491–536. 2010.

[14] C. E. Lathan and M. Tracey. The effects of operator spatial perception and sensory feedback on human-robot teleoperation performance. *Presence: Teleoper. Virtual Environ.*, 11(4):368–377, 2002.

[15] N. Mavridis, N. Giakoumidis, and E. Machado. A novel evaluation framework for teleoperation and a case study on natural human-arm-imitation through motion capture. *I. J. of Social Robotics*, 4:5–18, 2012.

[16] T. Röfer, T. Laue, J. Müller, A. Fabisch, F. Feldpausch, K. Gillmann, C. Graf, T. J. de Haas, A. Härtl, A. Humann, D. Honsel, P. Kastner, T. Kastner, C. Könemann, B. Markowsky, O. J. L. Riemann, and F. Wenk. B-human team report and code release 2011, 2011. http://www.b-human.de/downloads/bhuman11_coderelease.pdf.

[17] O. Schrempf, U. Hanebeck, A. Schmid, and H. Worn. A novel approach to proactive human-robot cooperation. In *RO-MAN*, 2005.

[18] M. Tenorth, D. Nyga, and M. Beetz. Understanding and executing instructions for everyday manipulation tasks from the world wide web. In *IEEE International Conference on Robotics and Automation*, 2010.