Research Article Open Access

Adeline Chanseau*, Kerstin Dautenhahn, Kheng Lee Koay, Michael L. Walters, Gabriella Lakatos, and Maha Salem

How does peoples' perception of control depend on the criticality of a task performed by a robot

https://doi.org/10.1515/pjbr-2019-0030
Received March 31, 2019; accepted September 18, 2019

Abstract: Robot companions are starting to become more common and people are becoming more familiar with devices such as Google Home, Alexa or Pepper, one must wonder what is the optimum way for people to control their devices? This paper presents an investigation into how much direct control people want to have of their robot companion and how dependent this is on the criticality of the tasks the robot performs. A live experiment was conducted in the University of Hertfordshire Robot House, with a robot companion performing four different type of tasks. The four tasks were: booking a doctor's appointment, helping the user to build a Lego character, doing a dance with the user, and carrying biscuits for the user. The selection of these tasks was based on our previous research to define tasks which were relatively high and low in criticality. The main goal of the study was to find what level of direct control over their robot participants have, and if this was dependent on the criticality of the task performed by the robot. Fifty people took part in the study, and each experienced every task in a random order. Overall, it was found that participants' perception of control was higher when the robot was performing a task in a semi-autonomous mode. However, for the task "carrying biscuits", although participants perceived to be more in control with the robot performing the task in a semi-autonomous mode, they actually preferred to have the robot performing the task automatically (where they felt less in control). The results also show that, for the task "booking a doctor's appointment", considered to be the most critical of all four tasks, participants did not prefer that the robot chose the date of the appointment as they felt infantilised.

Keywords: task criticality, perception of control, robot companion

***Corresponding Author: Adeline Chanseau:**
Adaptive Systems Research Group, University of Hertfordshire;
E-mail: a.chanseau@herts.ac.uk

1 Introduction

One of the first commercial domestic robots was the Roomba from IRobot [1]. Although this robot was single-task and its main purpose was to vacuum clean the house, researchers found that after having the robot for some time, the users tended to treat the robot as a pet, with affection [2]. Some companies used this finding to cleverly market their products. For example, Moulinex named one of its 2018 cooking robot range "robot - cuiseur companion" [3] meaning cooking robot companion. Other companies desperately tried to launch a domestic robot companion that can express some intelligence, with facial or voice recognition features, cameras or an advanced AI able to teach you Yoga. For example the Lynx robot by Ubtech [4]. Many of these devices although technically innovative, suffered from technological limitations in areas such as real time voice and gesture recognition, or the implementation of artificial intelligence to manage the house according to the owners habits and so on. However, it is likely that those limitations and challenges will be overcome in the near future. Therefore, researchers have speculated that having such a robot is ethically acceptable [5–7]. Bernotat and Eyssel [8] suggested that anxiety towards robots come from these unanswered ethical questions, in particular those that the popular culture sometimes portrays in the worst possible way.

In the late 1990s, Shneiderman and Maes [9] investigated the best way to control a computer. Later on, it turned out the mouse became the popular tool compared

Kerstin Dautenhahn: Adaptive Systems Research Group, University of Hertfordshire; E-mail: k.dautenhahn@herts.ac.uk
Kheng Lee Koay: Adaptive Systems Research Group, University of Hertfordshire; E-mail: k.l.koay@herts.ac.uk
Michael L. Walters: Adaptive Systems Research Group, University of Hertfordshire; E-mail: m.l.walters@herts.ac.uk
Gabriella Lakatos: Adaptive Systems Research Group, University of Hertfordshire; E-mail: g.lakatos@herts.ac.uk
Maha Salem: Adaptive Systems Research Group, University of Hertfordshire; E-mail: me@mahasalem.net

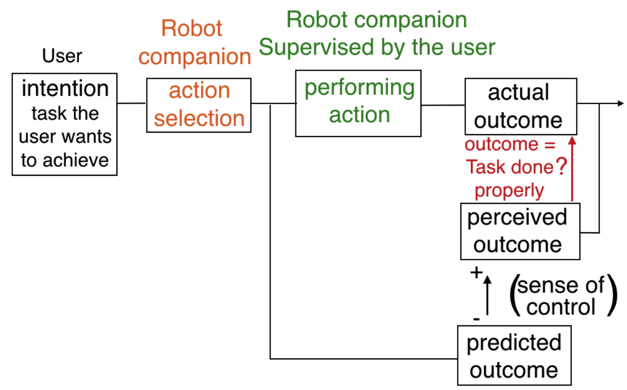

Figure 1: How perception of control is being studied in this paper [13].

to an agent [10]. With the development of robotics, the question of how to control a robotic system needs to be asked. Some previous research in psychology, links anxiety and perception of control [11]. Based on this and previous research performed in the research area of Human Computer Interaction (HCI) on user control, we conducted a preliminary live study [12] which demonstrated there is a link between anxiety towards robots and perception of control. However, this previous study had the robot performing only one task. This paper goes further by attempting to understand what level of control people want to have over their robot depending on the type of task the robot is performing. Are people ready for a robot that can take decisions on their behalf?

Our article presents an investigation into people's perception of control of their robot companion, by measuring how much supervision people consider the robot needs to perform its task correctly (Figure 1), or rather how much in control of the robot participants feel they need to be, depending on the task performed by the robot. Perception of control is a difficult topic to study as it needs to be explained clearly, what it is that we study. Based on Haggard and Chambon's schematic of sense of control [13], we adapted their schematic and simplified it for our current investigation (See Figure 1). When we consider an action, there are three steps which we think of: what is the outcome that we want, what is the action that we need to perform to reach this outcome, and how to perform the action to get it right. This is what Figure 1 displays. To be able to get an accurate measurement, the investigation is conducted with the robot being supervised by the user in one condition, and with the robot performing the action without supervision in the other condition. The measurement of perception of control of the robot companion is therefore conducted through the level of autonomy of the robot.

In one of our previous studies [12], we investigated what the preferred level of autonomy of the robot companion was, when the robot had to perform a cleaning task via a cleaning robot. So, the robot companion would either activate directly the cleaning robot with or without the acknowledgement of the user depending on the condition, or would send the user to activate the cleaning robot. We found that people preferred the more automated version of the robot companion, when the robot companion activated the cleaning robot without acknowledgement needed from the user. The results also showed that the more controlling a person's nature is, the more likely they will want to have a more autonomous robot.Note, by autonomous, we mean a robot can make decisions. To consolidate these results, it was chosen to conduct this current experiment with four different tasks. The tasks were chosen based on the results of a previous questionnaire study [14] which investigated what type of tasks are considered "high in criticality" or "low in criticality" from those categorised as mainly "physical" or "cognitive". We decided to use 2 highly critical tasks, one physical (carrying biscuits), and one cognitive (booking a doctor's appointment), and 2 low critical tasks, one physical (dancing), and one cognitive (building a Lego character), in order to balance the type of tasks for our current live experiment. Therefore this paper will study the preferred level of control of the robot depending on the criticality of the task performed by the robot.

The chosen tasks were also evaluated in our current live study to confirm the results of our questionnaire study. We tried to chose tasks that reflected what people might want to have performed in an everyday life situation. One of the tasks relates closely to a recent product launched by Google, Google Duplex. Google Duplex is an agent that possesses artificial intelligence which allows it to communicate verbally to people [15]. Google Duplex was not specifically tested, but the robot used in this experiment was able to book appointments in an automated way, which is what the Google product was seen to be capable of doing in its promotional video. The other tasks were typical everyday tasks, such as helping carrying objects (in this scenario carrying biscuits), or entertainment based tasks such as building a Lego or dancing.

2 Background research

2.1 How do we study perception of control?

To be able to study perception of control, it is important to understand what locus of control is. Lefcourt theorised it by explaining that locus of control is how much people believe they can affect the relationship between actions and outcomes [17]. Pacherie [16] explained there are three types of intentions we have before executing an action: the practical reasoning of how to perform the intended action (mental effort), the physical requirements to make the action possible (physical effort), and the specification of the movements that are needed to execute the action. To illustrate Pacherie's theory, we draw a simplified schematics of her action specification in Figure 2. Haggard and Chambon investigated the biological pattern of sense of control, which is the neurology of perception of control [13]. They identified sense of control as the difference between the perception of the outcome, and the predicted outcome. Therefore, to study the perception of control of an action performed by a robot, it is necessary to first identify the predicted outcome of the action performed by the robot. Another way to study perception of control as Pacherie mentioned [16] is also to identify the predicted execution of the action performed by the robot. There is then two ways to study perception of control, either by checking the outcome of the action and see if it matches the user's expectation, or either by verifying if the way the action is performed by the robot matches how the user expected it to be performed. For our current investigation, we decided to focus on the latter one.

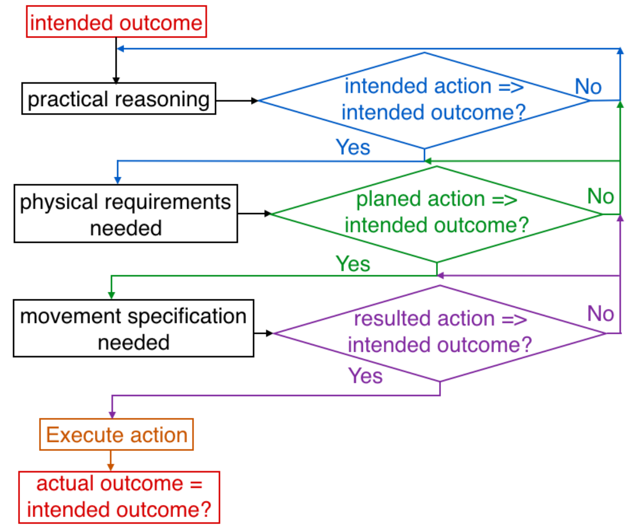

Figure 2: Simplified schematics of the mental process behind planning a specific action [16].

2.2 What is the difference between a task high in criticality and a task low in criticality?

Yanco and Drury were the first researchers to attempt of providing a clear definition of criticality in Human-Robot Interaction (HRI) [18]. They defined criticality as "the importance of getting the task done correctly in terms of its negative effects should problems occur" and a critical task as "to be one where a failure affects the life of a human". As Tzafestas later developed [19], we can distinguish three levels of criticality: high, medium and low. However, none of them specified how to quantify the failure, and neither how to measure its consequences on a human life. Guiochet, Machin and Waeselynck [20] studied safety critical robots. They mainly focused on industrial and advanced robots (robots that have decisional autonomy and are in a non-structured workplace), and detailed well the steps of how to evaluate a task. They looked at a task complexity, its function, and the type of safety rules that can be applied. However, thinking of a domestic companion, some household tasks that seemed simple such as ironing have proven to be complex to execute [21], as Dai, Taylor, Liu and Lin well explained in their paper. Ezer, Fisk, Rogers and Wendy [22] conducted a questionnaire study on a robot performing domestic tasks. They defined the criticality of a task depending on how much benefit participants perceived it to have. As such, tasks high in criticality were related to emergency tasks, and tasks low in criticality, to entertainment. However, they did not ask their participants to confirm their rating of criticality. This is why we conducted our previous study on task criticality [14], in order to identify what the users perceives as high critical or low critical tasks. We found that tasks related to entertainment were mostly rated low in criticality unless the robot is meant to be used only for this purpose. As our users underlined, they would consider a task high in criticality if the robot was mainly designed to do one particular task. For example, cleaning would be a task high in criticality for a Roomba. Our previous study showed that the tasks that people generally considered high in criticality were the ones that potentially had an irreversible effect or would have been difficult to change if a mistake was made (i.e. for example overcooking rice with the wrong quantity of water so it then becomes porridge).

Therefore, we decided to choose for this live investigation four tasks from the results of our questionnaire study. We wanted to vary the tasks, so we chose two tasks rated as high in criticality and two low in criticality. The definition of criticality used in this paper results from a previous study on task criticality [14]. As explained above, since the concept was difficult to grasp for participants, some statements about tasks were provided for them to rate. As a result, criticality can then be defined as **"the im-**

portance of a task being carried out safely, correctly and with attention to detail". To provide a good variation of tasks, it was also chosen to provide two cognitive tasks and two physical tasks according to participants' ratings from our previous study done on task criticality [14]. A cognitive task is defined as "any task that requires mental activities or thinking processes, which may involve some decision-making" and a physical task as "any task that requires body movement or motion processes, which may be qualified as a laborious task" [14].

As tasks high in criticality, we chose booking a doctor's appointment (cognitive), and carrying biscuits (physical) to implement in our scenario. Although carrying biscuits does not seem high in criticality, it was rated as such by participants because of its irreversibility aspect. If biscuits fall on the floor and break, there are no more biscuits to be eaten. This could be problematic when house owners are receiving guests and were planning to provide tea and biscuits. We chose as tasks low in criticality, building a Lego character (cognitive) as solving a puzzle was considered low criticality. We decided to modify the task into building a Lego model to make the task more interactive with the robot. The second low critical task we picked was doing a dance (physical), as it was a task that the robot could perform which was entertaining, which makes it low in criticality.

3 Research questions and hypotheses

According to our previous studies on perception of control and robot companions, and the literature, we formulated the following research questions and hypotheses for this paper:

- **R1**: Is there a relationship between participants' desire to control and their preference for the robot's level of autonomy?
  - H1: The more participants want to be in control, the less autonomous they want the robot to be.
    This hypothesis comes from the results of an early study on perception of control and robot companions [12].
- **R2**: Is there a relationship between the perception of control participants had over the robot and their preference of the robot's level of autonomy?
  - H2: The more in control over the robot participants perceive they are, the less autonomous they want the robot to be.
    This hypothesis comes from the results of the same study mentioned above on perception of control and robot companions [12].
- **R3**: Does the level of criticality (high or low) of a task performed by the robot influences participants' preferences for the robot's level of autonomy?
  - H3a: The higher the criticality of the task performed by the robot is, the less autonomous participants prefer the robot to be.
    The higher the criticality of a task is rated, the more risks it involves. Therefore, it can be hypothesised that the higher the criticality of a task becomes, the more participants prefer to control the robot, so the less autonomous they want the robot to be.
  - H3b: The more controlling a participant is, the higher they tend to rate the criticality of a task.
    We hypothesised that the higher the criticality of a task becomes, the less autonomous participants want to be, and therefore the more controlling participants are. Maybe the more controlling participants are, the more they rate tasks high in criticality.
- **R4**: Is there a relationship between the type of task performed by the robot (physical or cognitive) and participants' preference for the robot's level of autonomy?
  - H4: The participant's preference for the robot's level of autonomy is independent of the type of task performed by the robot.
    There is no evidence in the literature suggesting there is a connection between the robot's level of autonomy and the type of task the robot is performing. Hence the hypothesis rests on these two variables being independent.
- **R5**: Does a participant's tech savviness (experience and knowledge about technology) influence their preferences for the robot's level of autonomy?
  - H5: The participant's preference for the robot's level of autonomy is independent of their experience and knowledge about technolog.
    The result of our first study on perception of control [12] suggested that technology awareness is independent on the participant's preference of the robot's level of autonomy.

4 Method

To investigate how task criticality influences participant's choices for the level of autonomy of a robot companion, a live study was conducted in the Robot House (Figure 3).

This house is a typical British residential home owned by the University of Hertfordshire, and has been converted into a smart house, with the purpose to host HRI studies in a realistic domestic environment. We used a mobile robot for the experiment called Sunflower (Figure 4). As explained in the introduction, we previously conducted a live study on perception of control [12], and the measurement of the level of control was assumed to correspond to the level of autonomy of the robot. By autonomy, we mean decision-making. To remain consistent, we used the same type of measurement in this investigation, which led to two conditions: one in which the robot is supervised by the user while performing the action, and another in which the robot is performing the action without supervision from the user.

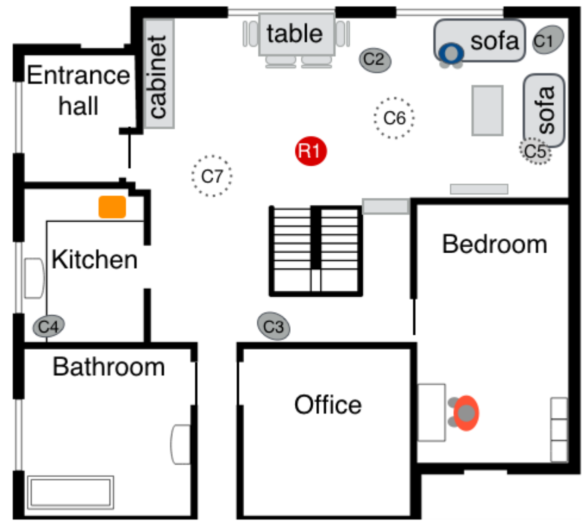

Figure 3: Setup of the experiment: the experimenter is in the bedroom able to see through live cameras C5, C6 and C7 what is happening in the house. The other cameras C1, C2, C3 and C4 were there just for recording purposes. R1 is the robot starting each interaction in a resting position.

4.1 Experimental design

To be able to investigate the influence of task criticality, four tasks were carefully selected for the experiment: a cognitive task low in criticality T1, a cognitive task high in criticality T2, a physical task low in criticality T3, and a physical task high in criticality T4. These tasks were classified and pre-validated in our previous questionnaire study [14]. Each task consisted of two conditions: one in which the robot was making decisions on how to perform the action, and the other one in which the robot was guided by the participant to perform the action. The tasks were performed by the Sunflower robot (see Figure 4), a custom-made robot that possesses a Pioneer DX robot base, a head and a tray. In the experiment, the robot navigation was autonomous, but the messages on its tablet, the tray movements, and the robot dancing movements were controlled by the experimenter.

4.1.1 Tasks and conditions of the experiment

- **T1 "Lego" cognitive task low in criticality**: The participant wants to build a Lego character with the help of Sunflower.
  - **C1 fully autonomous**: The robot decides when to show the next step of how to build the Lego character. The robot uses its tablet to show the next step in order to help the participant. The robot displays the next step as soon as it sees that the participant is finished.
  - **C2 semi-autonomous**: The participant decides when to see the next step of how to build the Lego character on the robot's tablet.
- **T2 "doctor" cognitive task high in criticality**: The participant has to do a blood test in the following days. Sunflower reminds the participant and offers to book the appointment.
  - **C1 fully autonomous**: The robot decides which appointment time slot to arrange for the doctor's appointment after checking the diary. The robot then confirms that a notification will be sent on the day of the appointment 2 hours beforehand. It is implicitly suggested that the robot put the appointment in the digital diary.
  - **C2 semi-autonomous**: The robot offers some appointment slots available and the participant chooses the one he/she prefers. The robot asks when the notification should be sent and offers options.
- **T3 "dance" physical task low in criticality**: The participant does a dance with Sunflower. To do so the user shows a dance movement to the robot. The robot then shows a dance movement to the participant. This is then repeated once.
  - **C1 fully autonomous**: The participant does one movement and Sunflower does another random movement to express creativity. For example, if the participant steped to the left, the robot would not move to its left but would step to another position. For example, forward).
  - **C2 semi-autonomous**: The participant does one movement and Sunflower repeats the movement

Figure 4: Sunflower robot.

(for example, if the participant turns right, Sunflower turns right). The same applies to all movements.

- **T4 "biscuits" physical task high in criticality**: The participant is expecting guests. Sunflower wants to help the participant to carry some biscuits for the guests to the living room.
  - **C1 fully autonomous**: As soon as Sunflower's tray is loaded, the robot goes to the living room.
  - **C2 semi-autonomous**: When the participant has finished loading Sunflower's tray, the participant provides voice commands to guide the robot to the living room by giving simple direction commands (go, left, right, stop, destination reached).

4.2 Participation

The experiment was conducted as a within-subjects design. Each participant experienced all task conditions of the experiment in a semi-randomised order (4x2). The experiment lasted between 45 minutes and 2h30 for the slowest participant. It took a mean time of 1h15 to complete the experiment.

Fifty participants (28 females and 22 males) were recruited from the University of Hertfordshire and its surroundings, using email advertisements and posters. They were tested individually. Each participant received five pounds sterling as a travel compensation to come to the Robot House. Their age range varied from 19 to 80 ($Mean$ = 39.98, SD = 14.88). Regarding technology awareness, every participant mentioned having a computer (86% of them use it daily, and 14% use it weekly). Ninety percent of our participants use their smartphone daily. The other ten percent did not possess a smartphone. Twelve percent interacted on a daily basis with either a Google Home or an Amazon Alexa. A five-point Likert scale questionnaire (1 being not familiar at all and 5 being very familiar) showed that our participants were mostly unfamiliar with programming robots ($Median$ = 1, SD = 1.06), had little experience programming robots ($Median$ = 1, SD = 0.91), and had little experience interacting with robots($Median$ = 1, SD = 1.03). Eighty-six percent of the participants had a job which was dominantly intellectual and cognitive (such as an office job, as an IT consultant or a lecturer). Eight percent of the participants had a more physical job such as being a golf professional or a bus driver. The rest of the participants mentioned being either retired or being a home-maker.

4.3 Experimental procedure

4.3.1 Greetings

Participants were asked to come directly to the Robot House for the experiment. Each of them were formally greeted and offered a tour of the Robot House. This allowed the experimenter to introduce the technology (the robot and sensors) and explain the purpose of the house. After this introduction, the visitor was given an information sheet, a consent form and an ID number (used for anonymisation purposes). A hot beverage was offered while forms were completed. Then the participant was asked to complete a questionnaire collecting data on demographics (age, gender, job...), technology experience, and familiarity with robots. This was followed by a Big Five personality test [23], which is well-used in HRI research, and the Desirability Control Scale (DCS) questionnaire [24] that has been used since the 1980s in psychology research to measure desired control. These questions were to help answer the research question R1, as the desired control level is measured by the DCS, and we then cross-correlate the data with answers form the personality data to see if there is a relationship between these variables. The questionnaire on technology savviness will help providing answers to the research question R5, as it allows a cross comparison between the frequency of the usage of some commonly-used technologies and the participants' preferences' for the robot's level of autonomy.

4.3.2 Introduction to Sunflower

Sunflower was then introduced to the participant as a robot companion that can help people. It was explained that the interaction with the robot would mainly happen in the living room and in the kitchen. After this, the participant was asked to sit on the sofa and was given a set of questions assessing the participant's expectations of the Sunflower robot. This allowed the experimenter to prepare the robot for the first interaction session and to turn on the cameras. One of the four scenarios was presented to the participant. He or she was told that the same scenario would occur twice in a row. The experiment was designed this way so that after each scenario, participants could do an immediate comparison on the two conditions (C1 the Sunflower robot being fully autonomous, or C2 the robot being semi-autonomous). So, each participant could experience each task in a semi-randomised order (4x2). The randomisation was counterbalanced, as half of the participants started the experiment with the first condition C1, and the other half started with the second condition C2. A quarter of the participants started the experiment with Task 1, a quarter with Task 2, a quarter with Task 3 and a quarter with Task 4.

4.3.3 Interaction phase

Once the robot was setup, the experimenter told the participant they can interact with the robot as soon as the experimenter left.

Task 1 "You have some time off and want to build a Lego character with the robot."

Condition 1 Sunflower came to the participant and displayed the following message to the participant on its screen: "Today we are going to build a Lego character together. I will guide you through the process. Please once you are ready, say 'ready' so we can start". The participant had to click or say 'ready' to start the process. Then the Sunflower robot opened its tray to deliver the Lego pieces and started displaying the image instructions on how to build the Ironman Lego character. As soon as the participant finished the first step, the robot showed the next step. Once the Lego was built, the experimenter came out of the room and said to the participant "Now the same scenario will start again". The experimenter provided new pieces of Lego on the robot's tray.

Condition 2 The same process started again except that this time, the robot mentioned that the participant had to say 'next' to see the next instruction page. Once the session was over, the experimenter came out of the room and provided a set of questionnaires to the participant which contained questions on their perception of control, criticality and habituation to building Lego available at [25].

Task 2 "You have just come back from a trip to Indonesia and you need to do a blood test to check for Dengue fever in the following days."

Condition 1 Sunflower came to the participant and displayed the following message on its screen: "You need to do your blood test soon. Let me check your diary to see when you are available next for a blood test. I will check with the NHS when your appointment can be booked for." A waiting message appeared next "Checking....". The robot then said: "I have found a free slot for you. I have added it to your digital calendar. I will send you a reminder the evening before the appointment and a notification 2 hours before the appointment." The experimenter came out of the room and said: "Thank you. Now the same scenario will start again."

Condition 2 Sunflower came to the participant and displayed the following message: "You need to do your blood test soon. Let's check on your diary when you are available next for a blood test. I will check with the NHS when they have a free slot." But this time the robot offered free slots in a calendar format for the user to choose from: "I have found these slots for you. Please pick the one you prefer." Once the participant made a choice, the robot offered to choose when the notification of the appointment should be made: "Thank you, your appointment has been booked. When shall I give you a reminder?" The robot offered several options. Once the participant chose an option, the robot said "Thank you, your choice has been recorded." Once the session was over, the experimenter came out of the room and provided a set of questionnaires to the participant. The questionnaire asked about perception of control, criticality and digital calendar [26].

Task 3 "You want to do a dance with the Sunflower robot and show some movements. You will show a sequence of 2 movements from the list below, in any order you like, one step at a time. The list is: move right, move left, move forward, move backward. You can also say it out loud to help the robot identifying the movement."

Condition 1 Sunflower came to the sofa and offered to the user to do a dancing activity together. The robot positioned itself and waited for the participant to start. After each movement that the participant did, the robot produced a random movement different from the one shown by the participant. Once the dance was over, the experimenter came out of the room to mention that the same scenario would start again.

Condition 2 The same routine happened except that the robot repeated each movement the participant made. Once the two dance steps were done, the participant was given another set of questionnaires containing questions about their perception of control, criticality and entertainment [27].

Task 4 "You are about to receive some guests home. You need some help from Sunflower to carry biscuits from the kitchen to the living room."

Condition 1 As in every scenario, the participant sat on the sofa and the robot came to them. Sunflower reminded the user that some guests were coming: "Hello, you are about to receive guests, let me help you to carry some biscuits from the kitchen to the living room. Let's go to the kitchen." Then the robot and the participants went to the kitchen. Once there, Sunflower opened its tray: "My tray is open. Please put one biscuits box inside." As soon as the participant loaded the robot's tray with a biscuit box (three biscuits boxes are on display in the kitchen), Sunflower went to the living room. Once in the living room, the robot asked for the tray to be unloaded: "Please take the biscuits box off my tray." Once it was done it displayed a thank you message: "I hope I was useful. I was happy to help you :)" The experimenter then came in the room and reset the scenario.

Condition 2 The following scenario was given to the participant: "You are about to receive some guests home. You need some help from Sunflower to carry biscuits from the kitchen to the living room. To guide the Sunflower robot, you can give the following commands: go, stop, left, right, destination reached." The same process started again but once the robot reached the kitchen, it reminded the participant it needed to be guided back: "Please guide me with the following commands: go, stop, left, right, destination reached." The participant then said a command and the robot responded. As soon as the living room was reached, the robot displayed another thank you message. The experimenter then provided another set of questionnaire containing questions on perception of control and criticality [28].

4.3.4 Last questionnaire and the reward

After the interaction phase, the user was given one last set of questionnaires that gained their perception of the criticality of each task and to provide some information on the overall interaction. The participant was then offered to take a "selfie" with the Sunflower robot displaying a personalised message "Hello *name*, it was nice to meet you :)" and £5 was given as a travel compensation.

4.4 Statistical analysis

As a lot of data was collected for this experiment, the data analysis was systematically carried out in this order:

- a descriptive analysis was performed to assess the general trends of the dataset.
- a Kolmogorov-Smirnov normality test was applied to see what type of correlation test can be performed.
- the dataset that was to be presented are non-parametric. Therefore, a Kendall's tau correlation test was used.
- when a correlation test was not possible to be used due to categorical nominative data, a Pearson Chi Square test was used to measure associations.

Correlation tests were chosen for this study as we wanted to quantify the association between variables, instead of comparing groups.

5 Results

5.1 Preferred conditions for each task

To evaluate which conditions people preferred for each task, participants answered a multiple choice questionnaire and had to provide the reason for their choice. As the pie charts Figure 5 show, there is clearly a preference for the semi-autonomous condition, when the user tells Sunflower how to perform the task; for T1, building a Lego model, T2, booking a doctor's appointment and T3, dancing. It can be noted that less than the majority preferred the semi-autonomous condition C2 for task 3. Sixteen percent of the participants preferred the autonomous condition C1 for T4 because they "retained control", one participant even said that he "wanted to have the biscuit on the other table, not where [he] was originally sitting", while another 6% said that having both options would be good to "test how much Sunflower is reliable" in the C2

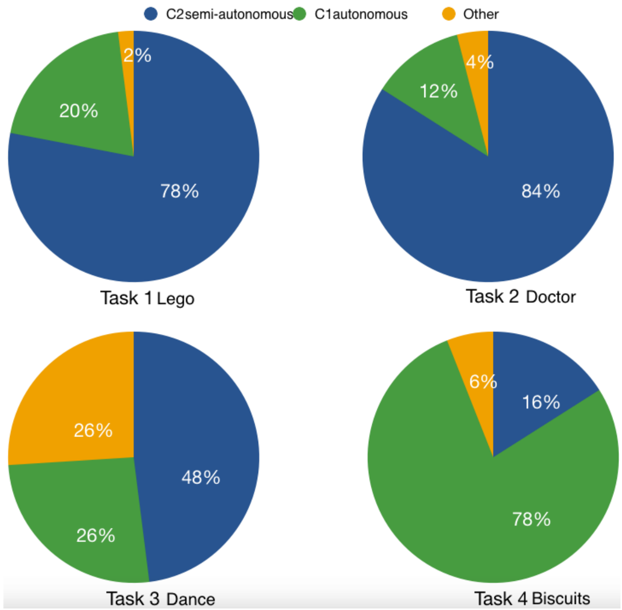

Figure 5: Preferred conditions for each task.

	Preferred condition Task 1	Preferred condition Task 2	Preferred condition Task 3	Preferred condition Task 4
Desired control	-0.129 0.273	-0.113 0.335	-0.047 0.676	-0.097 0.408
Preferred condition Task 1		-0.010 0.943	0.317* 0.017	0.022 0.876
Preferred condition Task 2			-0.080 0.549	0.088 0.522
Preferred condition Task 3				-0.297* 0.025

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 6: Correlation between desired control and the preferred conditions for each tasks.

semi-autonomous condition before fully adopting it. The majority, 78% of the participants preferred the C2 semi-autonomous condition for T4.

The Kendall's tau correlation test Figure 6 shows that there is no correlation between how controlling people are and their preferred condition for each task. However, there is a significant positive correlation between the preferred condition for Task 1 and the preferred condition for Task 3 ($\tau_b = 0.317, p = 0.017$). So, the more participants preferred the robot to be controlled in T1, the more they preferred the robot to be controlled in T3. There is a significant negative correlation between the preferred condition for Task 3 and the preferred condition for Task 4 ($\tau_b = -0.297$, $p = 0.025$). So the more people preferred the robot to be controlled for Task 3, the more autonomous they wanted it to be for Task 4. This demonstrates consistency between the preferred choice of condition for tasks low in criticality.

To conclude it seems that participants prefer to be in control of the robot unless it is less efficient for the task to be done.

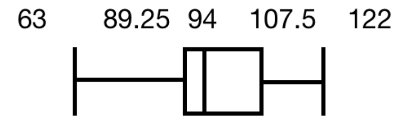

Figure 7: Desirability Control Scale.

5.2 R1. Personality effect

To measure how controlling people are, we used the DSC [29]. This test uses everyday life questions to study how much in control people want to be in general. A Kolmogorov-Smirnov normality test revealed that we had a normal distribution population in terms of how controlling people are ($D(50) = 0.104, p = 0.200$), which is illustrated in the boxplot Figure 7 and the histogram in Figure 8. The histogram shows that although the experiment had a good number of participants for a HRI live study ($N = 50$), the graph clearly shows there is not a proper normal distribution. This means that all the following statistical tests used were non-parametric, and therefore not strong enough to make definitive conclusions. However, the results still provided us with valuable information on tendencies on how people prefer to control their robot.

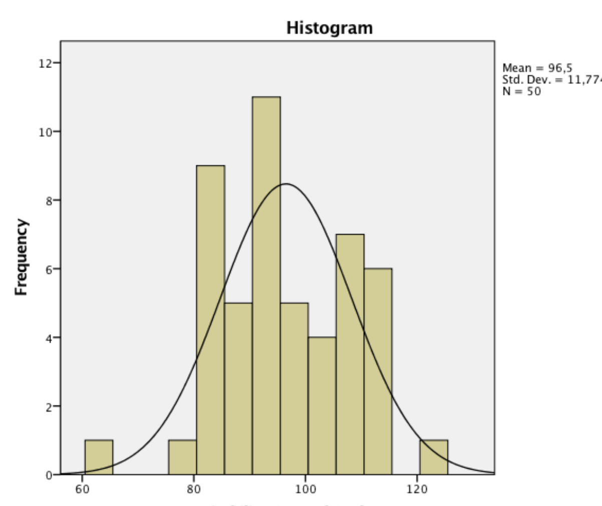

Figure 8: Desirability Control Scale population distribution.

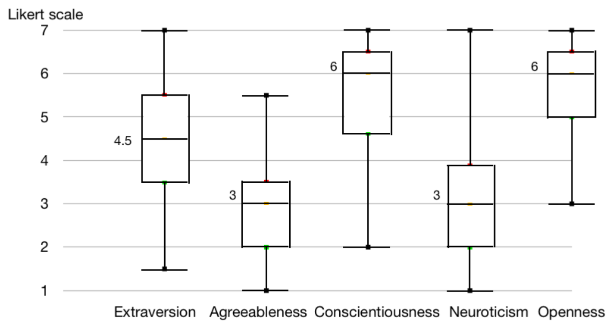

Figure 9: Boxplot of the Big Five personality test results.

To measure their personalities, participants were asked to respond to the standard Big Five personality test [23]. This test is a 7 point Likert scale that measures extraversion, agreeableness, conscientiousness, neuroticism sometimes called emotional stability and openness. The results show that our participants were open-minded with a median score of 6 ($SD = 1.41$) and conscientious with a median score of 6 ($SD = 1.22$). The median score of neuroticism was high with a score of 6 ($SD = 1.51$), while the median score of agreeableness and extraversion were low (3,$SD = 1.22$ and 4.5, $SD = 0.92$ respectively). As the boxplots of the personality test displays in Figure 9, our pool of participants do not represent a normal distribution. The Kolmogorov-Smirnov normality test confirmed that Agreeableness ($D(50) = 0.144, p = 0.011$), Conscientiousness ($D(50) = 0.220, p < 0.001$), Neuroticism ($D(50) = 0.126, p = 0.045$) and Openness ($D(50) = 0.155, p = 0.004$) do not follow a normal distribution. Therefore it was decided to use non-parametric correlation tests. A Kendall's tau correlation test indicated that the more open-minded people were, the more controlling they tended to be, through a significant positive correlation test between openness and the desirability control scale ($\tau_b = 0.254, p = 0.016$).

There is a highly significant negative correlation between Openness and the preferred condition participants had for Task 1, building a Lego character ($\tau_b = -0.323, p = 0.009$). This means that the more openminded participants were, the less they preferred the semi-autonomous version of the robot (when they indicated when Sunflower had to show them the next step of the instructions to build the Lego character), therefore the more willing participants were to have a fully autonomous version of the robot (when Sunflower decided when to show the user the next step of the instructions.). For the other tasks, there is no statistically significant correlation between Openness and the preferred condition participants had for Task 2 ($\tau_b = -0.126, p = 0.309$), Openness and the preferred condition participants had for Task 3 ($\tau_b = -0.177, p = 0.138$), and Openness and the preferred condition participants had for Task 4 ($\tau_b = -0.26, p = 0.832$).

A Kendall's correlation test showed there is no statistically significant correlation between personality traits and the perception of control in the participants' preferred condition for Task 1. However, there is a statistically significant negative correlation between the Desirability Control Scale ratings and the perception of control of the Sunflower robot for Task 2 ($\tau_b = 0.274, p = 0.014$). So the more controlling the participant is, the less he/she felt in control of the robot when the doctor appointment was booked. There was no significant correlation between personality traits and the perception of control in the user's preferred condition for Tasks 3 and 4. Therefore, no generalisation can be made regarding how controlling people are and how autonomous they want the robot to be.

5.3 R2. Perception of control of the robot companion

Participants were asked to rate on a scale from 1 to 5, (1 being "I didn't feel in control at all", and 5 being "I felt I was fully in control"), how much they felt in control of the action, how much they felt in control of the outcome of the action and how much they felt in control of the robot during the task, for both conditions (C1: when the robot decides what to do next, and C2: when the participant decides what the robot does next). First, as expected, the results show that people felt more in control when they decided what the robot had to do for each task (see Figure 10 for Task 1, Figure 12 for Task 2, Figure 14 for Task 3 and Figure 16 for Task 4). The results are consistent across each tasks. When the user did not feel in control of the action, the user also did not feel in control of the robot (see Figure 10 for Task 1 for example).

5.3.1 Task 1: Building a Lego Character

The results showed that on average, participants preferred the C2 semi-autonomous condition compared to the C1 autonomous condition for Task 1 "Lego": 78% of the participants preferred the C2 semi-autonomous condition when the robot displayed the following instruction after the user asked for it , while only 20% chose the C1 autonomous condition when the robot chose when it was appropriate to display the following instruction, as their preferred choice. 2% was undecided.

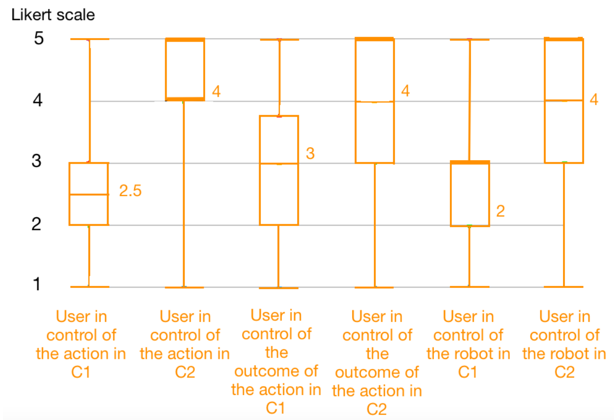

Figure 10: Task 1: building a Lego character.

	User in control of the action in the other condition	User in control of the outcome of the action in the preferred condition	User in control of the outcome of the action in the other condition	User in control of the robot in the preferred condition	User in control of the robot in the other condition	User's preferred condition in Task 1
User in control of the action in the preferred condition	-0.230 0.065	0.657** 0.000	-0.140 0.261	0.546** 0.000	-0.144 0.248	-0.073 0.587
User in control of the action in the other condition		0.165 0.183	0.591** 0.000	-0.237 0.053	0.500** 0.000	-0.158 0.233
User in control of the outcome of the action in the preferred condition			-0.076 0.539	0.689** 0.000	-0.001 0.992	0.111 0.408
User in control of the outcome of the action in the other condition				-0.174 0.156	0.594** 0.000	-0.051 0.702
User in control of the robot in the preferred condition					-0.139 0.258	0.186 0.162
User in control of the robot in the other condition						-0.014 0.916

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 11: Correlation table for Task 1 "Lego" between the participant's preferred condition and the participant's perception of control for their preferred condition and their non-preferred condition.

The Kendall's tau correlation test Figure 11 shows there is statistically a highly significant positive correlation between the perception of control of the action, and the perception of control of the outcome of the action for both conditions (for the preferred condition $\tau_b = 0.657$, and $p < 0.001$, for the other condition ($\tau_b = 0.591, p < 0.001$). There is also a highly significant positive correlation between the perception of control of the outcome of the action and the perception of control of the robot for both conditions (for the preferred condition $\tau_b = 0.689, p < 0.001$, for the other condition $\tau_b = 0.594, p < 0.001$). This means that the more people felt in control of the action building the Lego character, the more they felt in control of the outcome of the action (having a Lego character built) and the more they felt in control of the robot. This is true for the user's preferred condition and the user's non-preferred condition, which demonstrate the consistency of the results for Task 1.

5.3.2 Task 2: Booking a doctor's appointment

For Task 2, the descriptive statistics Figure 12 tell us there is an even clearer difference for perception of control between C1 (the robot being fully autonomous) and C2 (the robot being semi-autonomous) compared to Task 1 Lego (Figure 10). Participants clearly did not feel in control at all in the scenario ("booking a doctor's appointment") in C1, when the robot chose the appointment time slot for its user. For this Task 2, 84% of the participants preferred the C2 semi-autonomous condition when they decided their time slot for the doctor's appointment, while 12% preferred the C1 condition, when the robot chose the time slot, and 4% were undecided.

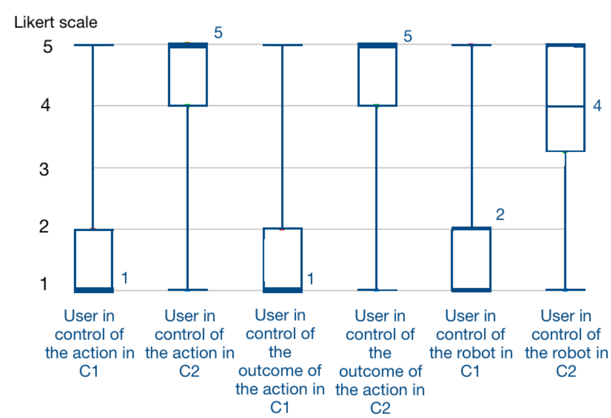

Figure 12: Task 2: booking a doctor's appointment.

The Kendall's tau correlation Figure 13 test reveals the same consistency found for Task 1: there is a statistically strong positive correlation between the perception of control of the action and the perception of control of the outcome of the action for both conditions (for the preferred condition$\tau_b = 0.805, p < 0.001$, and for the other condition $\tau_b = 0.701, p < 0.001$). There is also a strong significant negative correlation between the perception of control of the action for the preferred condition and the perception of control of the action for the other condition ($\tau_b = -0.498, p < 0.001$). This means that the more in control of the action "booking a doctor appointment" the participant perceived to be, the more in control of the outcome (doctor's appointment booked) the participant felt and the more in control of the robot the participant perceived to be in his/her preferred condition, in this case mostly C2 semi-autonomous (see Figure 12), it also means that the more in control of the action the user felt in his/her

	User in control of the action in the other condition	User in control of the outcome of the action in the preferred condition	User in control of the outcome of the action in the other condition	User in control of the robot in the preferred condition	User in control of the robot in the other condition	User's preferred condition in Task 2
User in control of the action in the preferred condition	-0.498** 0.000	0.805** 0.000	-0.453** 0.000	0.701** 0.000	-0.463** 0.000	-0.289* 0.037
User in control of the action in the other condition		-0.463** 0.000	0.868** 0.000	-0.282* 0.027	0.749** 0.000	0.193 0.162
User in control of the outcome of the action in the preferred condition			-0.494** 0.000	0.698** 0.000	-0.471** 0.000	-0.165 0.234
User in control of the outcome of the action in the other condition				-0.407** 0.001	-0.787** 0.000	0.183 0.179
User in control of the robot in the preferred condition					-0.394** 0.001	-0.177 0.190
User in control of the robot in the other condition						0.211 0.117

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Figure 13: Correlation table for Task 2 "Doctor" between the participant's preferred condition and the participant's perception of control for their preferred condition and their non-preferred condition.

preferred condition, the less in control the user felt in the other condition.

5.3.3 Task 3: Doing a dance

Task 3 descriptive statistics results Figure 14 were less pronounced than the ones for Task 1 and for Task 2, in terms of differences between the C1 fully autonomous condition and the C2 semi-autonomous condition. However, the results still show that people perceive to be more in control in the C2 semi-autonomous condition compared to the C1 fully autonomous condition. 48% of the participants preferred the C2 condition when the robot was repeating the user's dance step, while 26% of the users preferred the C1 condition when the robot was doing an unpredictable dance step and 26% were not sure what they prefer. The Kendall's tau correlation test Figure 15 shows the same statistical significant correlation seen for Task 1. There is a highly significant positive correlation between the perception of control of the action "doing a dance" and the perception of control of the outcome of the action for both conditions (for the preferred condition $\tau_b = 0.887, p < 0.001$, and for the other condition $\tau_b = 0.793, p < 0.001$). We found the same consistency in the correlation results for Task 3 contrasted to the ones for Task 1. So the more the user perceived to be in control of the action "doing a dance", the more the user felt in control of the outcome of the action (dance done) and the more the user perceived to be in control of the robot in the preferred condition. The same results were found in the other condition.

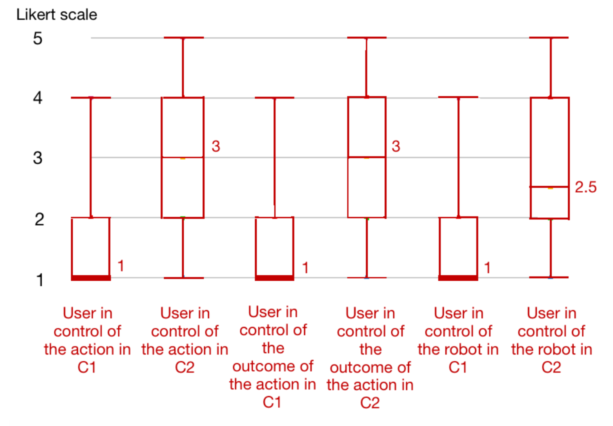

Figure 14: Task 3: doing a dance.

	User in control of the action in the other condition	User in control of the outcome of the action in the preferred condition	User in control of the outcome of the action in the other condition	User in control of the robot in the preferred condition	User in control of the robot in the other condition	User's preferred condition in Task 3
User in control of the action in the preferred condition	-0.126 0.378	0.887** 0.000	-0.104 0.468	0.753** 0.000	-0.014 0.925	0.066 0.651
User in control of the action in the other condition		-0.73 0.613	0.793** 0.000	0.052 0.718	0.798** 0.000	0.154 0.304
User in control of the outcome of the action in the preferred condition			-0.085 0.556	0.833** 0.000	0.055 0.706	0.080 0.583
User in control of the outcome of the action in the other condition				-0.007 0.963	0.795** 0.000	0.147 0.329
User in control of the robot in the preferred condition					0.202 0.164	0.278 0.57
User in control of the robot in the other condition						0.227 0.136

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 15: Correlation table for Task 3 "Dance" between the participant's preferred condition and the participant's perception of control for their preferred condition and their non-preferred condition.

5.3.4 Task 4: Carrying biscuits

It can be noticed that Task 4 descriptive statistics Figure 16 displays less differences between C1, the fully autonomous condition, and C2, the semi-autonomous condition, as compared to the other tasks. For Task 4, 18% of the users preferred C2 when they guided the robot to the living room, while 78% of them preferred C1 when the robot decided to go to the living room on its own, and 6% were undecided. The Kendall's tau correlation test Figure 17 indicates a statistically strong significant positive correlation between the perception of control of the action and the perception of control of the outcome of the action for both conditions($\tau_b = 0.671, p < 0.001$ for the preferred condition, and $\tau_b = 0.776, p < 0.001$). There is also a highly significant positive correlation between the perception of

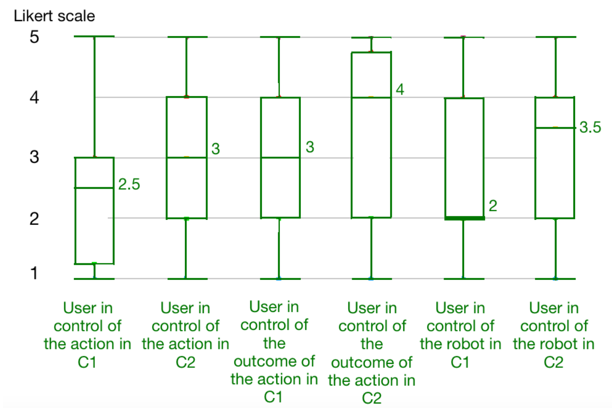

Figure 16: Task 4: carrying biscuits.

	User in control of the action in the other condition	User in control of the outcome of the action in the preferred condition	User in control of the outcome of the action in the other condition	User in control of the robot in the preferred condition	User in control of the robot in the other condition	User's preferred condition in Task 4
User in control of the action in the preferred condition	-0.090 0.462	0.671** 0.000	-0.139 0.258	0.718** 0.000	-0.168 0.171	0.127 0.343
User in control of the action in the other condition		-0.071 0.557	0.776** 0.000	-0.093 0.453	0.780** 0.000	0.116 0.377
User in control of the outcome of the action in the preferred condition			-0.099 0.415	0.641** 0.000	-0.178 0.143	0.043 0.747
User in control of the outcome of the action in the other condition					0.803** 0.000	-0.046 0.728
User in control of the robot in the preferred condition						-0.017 0.901
User in control of the robot in the other condition						-0.039 0.768

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 17: Correlation table for Task 4 "Biscuits" between the participant's preferred condition and the participant's perception of control for their preferred condition and their non-preferred condition.

control of the outcome of the action and the perception of control of the robot for both conditions (for the preferred condition $\tau_b = 0.641, p =< 0.001$, and for the other condition $\tau_b = 0.803, p < 0.001$). This means that the less the user perceived to be in control of the action "carrying biscuits", the less the user felt in control of the outcome "biscuits carried to the living room", and the less he/she perceived they were in control of the robot in his/her preferred condition. This is really interesting. For the majority of participants, the preferred condition was C1 when the robot decided to carry the biscuits to the living room on its own without guidance from the participant.

To conclude, we can see that the results are consistent across the tasks and conditions. When people felt in control of the robot for one condition, they would also feel in control of the action and in control of the outcome of the action for this same condition. These results did not validate the H2 hypothesis. There is a lot more subtlety into the link between how the user perceives to be control of the action/outcome of the action/robot and the participant's preferred choice of level of autonomy of for the robot. The results show that the more in control the participants felt for Task 2, Doctor, and Task 4, Biscuits, the more autonomous they wanted the robot to be. However, for Task 3, Dance, it was found that the more the participant felt in control, the less autonomous they wanted the robot to be.

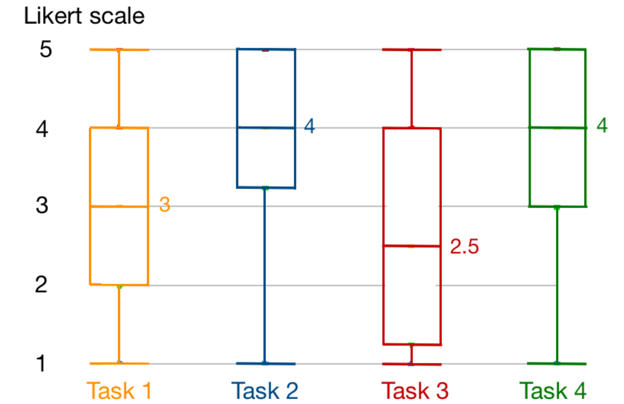

Figure 18: Task criticality results.

5.4 R3. Task criticality

To validate the classification of task criticality, participants were asked to rate the criticality of the task performed by the robot from a scale of 1 to 5, (1 being low critical and 5 being high critical). As observed on Figure 19, Task 2 Doctor and Task 4 Biscuits were rated as highly critical tasks ($Median_{task2} = 4$ $SD_{task2} = 1.11$ and $Median_{task4} = 4$, $SD_{task4} = 1.25$). Task 1 Lego seemed to be rated medium critical ($Median_{task1} = 3$, $SD_{task1} = 1.35$) and Task 3 Dance as low critical ($Median_{task3} = 2.5$, $SD_{task3} = 1.32$).

At the end of the experiment, participants were asked to rank the tasks between them from the most critical one to the least. Booking a doctor appointment was considered the most critical task by 72% of the participants. The second most critical task was carrying biscuits with 58%. It is to be noted that 18% considered this task the most critical task. The third most critical task was building a Lego character with 50% (26% ranked this task as the least critical), and finally dancing was considered the least critical task by 64% of the participants.

The Kendall's tau correlation test indicated a statistically highly significant positive correlation between users' rating of criticality of Task 1 Lego, and users' rating of crit-

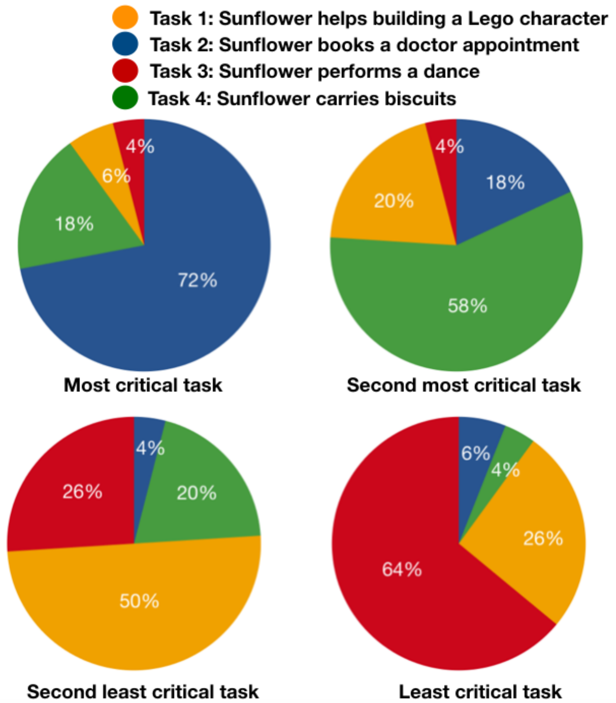

Figure 19: Task criticality rankings.

icality of Task 3 Dance ($\tau_b = 0.366, p = 0.002$), and between users' rating of criticality of Task 1 Lego and users' rating of Task 4 Biscuits($\tau_b = 0.356, p = 0.002$). This means, the more critical people considered Task 1 Lego to be, the more critical they would consider Task 3 Dance and Task 4 Biscuits to be too. There is also a strong significant positive correlation between people's criticality rating of Task 2 Doctor and people's criticality rating of Task 4 Biscuits ($\tau_b = 0.469, p < 0.001$). So the more critical people rated Task 2 Doctor, the more critical they also rated Task 4 Biscuits which is consistent with the way people ranked tasks among them, as Task 2 Doctor and Task 4 Biscuits were rated as the most critical tasks Figure 19, and the same goes between Task 1 Lego and Task 3 Dance as they were both rated the least critical tasks.

The Kendall's correlation test also showed there is a high significant positive correlation between participants' Task 3 criticality ratings and their Task 4 criticality ratings ($\tau_b = 0.388, p = 0.001$), which means the more critical people thought Task 3 was, the more critical they thought Task 4 was too. It could be explained by the way Task 3 criticality was rated, as the boxplot in Figure 19 displays a much wider spread rates compared to the other tasks. This is probably due to the way participants interpreted Task 3, doing a dance with the robots, as some may have thought they had to teach the robot, and this might have increased the criticality of this task.

However, the test revealed no statistically significant correlations between the way people rated the criticality of a task and their choice of preferred condition for any tasks. Therefore hypotheses H3a and H3b were not verified. We cannot say that the level of criticality of a task correlates with the choice of level of control of the Sunflower robot. Also, there was no correlation between the way people rated the criticality of a task and their desired control. It means that it cannot be said that the more controlling people are, the more critically they tend to rate tasks. Therefore H3c was not verified either.

5.4.1 Low critical tasks and perception of control

The test showed no correlations between participants' Task 1 criticality ratings and their perceptions of control of the action, their perceptions of control of the outcome of the action, or the perception of control of the robot, for any of the tasks and conditions. This means that the way participants rated Task 1, building a Lego character as a low critical task, did not influence the way participants perceived to be in control of the action executed in any of the 4 tasks, or the way participants felt in control of the robot during the performance of any of the 4 tasks, or the way participants perceived to be in control of the outcome of the action in any of the 4 tasks, in either conditions. The same results were displayed for Task 3 criticality ratings and the participants' perceptions of control of the action/outcome of the action/robot for any of the tasks and conditions.

5.4.2 High critical tasks and perception of control

5.4.2.1 Task 2: Booking a doctor's appointment

Task 2 Doctor criticality rating are statistically significantly positively correlated with the perception of control of the robot in T1/C2 Lego semi-auto when the participant decided when the robot displayed the next instruction ($\tau_b = 0.255, p = 0.037$). This means, the more people rated Task 2, as a high critical task, the more they felt in control of the robot in Task 1 when they decided when the robot displayed the next instructions on its tablet to build the Lego character.

There is a positive correlation between Task 2 Doctor criticality rating, and the perception of control of the action in T2/C2 Doctor/semi-auto when the participants decided when to book the doctor's appointment ($\tau_b = 0.394, p = 0.002$), and the perception of control of the outcome of the action in T2/C2 Doctor/semi-auto ($\tau_b =$

$0.321, p = 0.011$), and the perception of control of the robot in T2/C2 Doctor semi-auto($\tau_b = 0.303, p = 0.013$). So the more critical people thought Task 2 was, the more in control of the action "booking a doctor appointment" they felt, when they chose the time slot to be booked. They also felt more in control of the outcome, as they picked the time slot of the appointment, and they felt more in control of the robot, as the robot was following the user's instructions. This result shows consistency within Task 2 perception of control.

There is also a positive correlation between participants' Task 2 Doctor criticality ratings and their perceptions of control of the action in T4/C1 Biscuits/auto, when the robot decided to carry the biscuits to the living room on its own ($\tau_b = 0.354, p = 0.003$). So the more critically Task 2 Doctor was rated, the more people felt in control of the action "carrying biscuits" when the robot was manoeuvring without instructions.

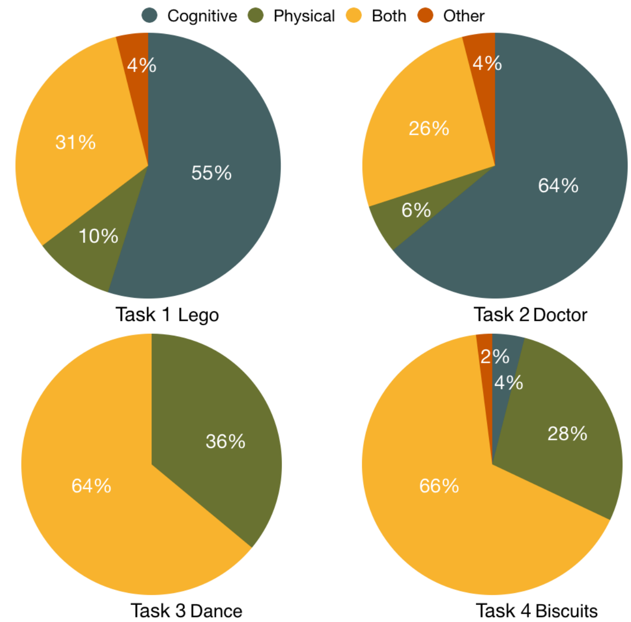

Figure 20: Classification of tasks according to their type.

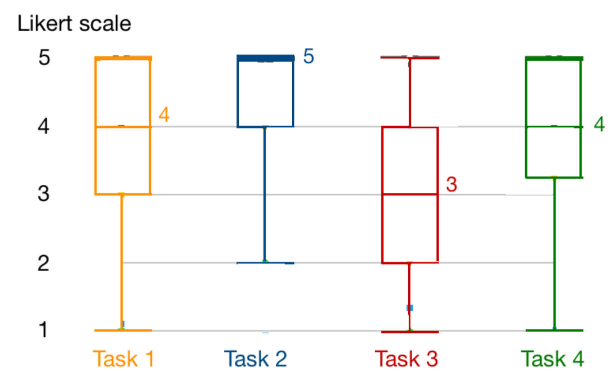

Figure 21: Boxplots of how realistic participants rated tasks according to a 5-point Likert Scale, 1 being not realistic at all and 5 being very realistic.

5.4.2.2 Task 4: Carrying biscuits

The test revealed a statistically significant positive correlation between Task 4 criticality ratings and perceptions of control of the robot in T1/C2 Lego/semi-auto ($\tau_b = 0.274, p = 0.022$). This means that the more critically people rated Task 4, the more in control of the robot they felt when they chose when Sunflower displayed the next step of the instructions to build the Lego character.

There is a significant positive correlation between Task 4 criticality ratings and perceptions of control of the outcome of the action in T2/C2 Doctor/semi-auto ($\tau_b = 0.285, p = 0.022$). So the more critical Task 4 is for a participant, the more in control of the outcome (appointment booked) the participant feels in Task 2, when the robot was following the instructions of the user.

There is also a significant positive correlation between Task 4 criticality ratings and the perceptions of control of the outcome of the action in T4/C2 Biscuits/semi-auto when the participant guided the robot to the kitchen with vocal commands ($\tau_b = 0.244, p = 0.037$). So the more critically participants thought Task 4 was, the more in control of the robot they perceived when they were guiding the robot from the kitchen to the living room. This result is interesting as participants did not actually prefer this condition for this task. So although participants felt more in control when guiding the robot, they still preferred the autonomous condition C1 for this task(when the robot chose to go to the living room as soon as the tray was full). It is most probably because participants prefer efficiency for this physical task as compared to control. As many stated, "the [robot] autonomous movement was faster and it required less effort" from the participant.

To conclude, these results demonstrate there are statistically significant correlations between the rating of task criticality and the perception of control of the robot when the task is considered critical. When the task is not considered critical, conclusions cannot be drawn regarding the importance of perception of control of the action/outcome of the action/robot. So hypotheses H3a and H3b were not verified.

5.5 R4. Type of task

Participants were asked to classify tasks according to their type (cognitive, physical, both or other). As expected, the choice of tasks was validated by the participants classification. Task 1 Lego and Task 2 Doctor were considered cognitive by respectively 55% and 64% of the participants. Task 3 Dance was either considered physical or both cognitive and physical by respectively 36% and 64% of the users. Task 4 Biscuits was classified as physical by 28% of the people and both by 66%. A Pearson Chi square test indicated a statistically significant association between the way participants classified Task 2 Doctor and their choice of preferred condition for Task 1 Lego ($df(6) = 15.783, p = 0.015$). There is also a significant association between users' classifications of Task 2 and their choices for preferred conditions for Task 3 Dance($df(6) = 14.158, p = 0.028$). And finally there is a significant association between people's classification of Task 4 Biscuits and their choice of preferred condition for Task 4 Biscuits($df(6) = 16.873, p = 0.010$). So the type of task seem to influence participants' choice of preferred condition which disproves hypothesis H4, but provides an explanation as to why the Task 4 Biscuits preferred condition was the autonomous condition C1, even though participants felt less in control. It is more probably because the task was physical, therefore participants may have found it more tedious to micromanage than a cognitive task.

People were asked to rate from the scale of 1 to 5 (1 being not realistic at all, and 5 being very realistic), how realistic they thought the task was. Task 2 Doctor and Task 4 Biscuits were rated as very realistic ($Median_{task2} = 5$, $SD_{task2} = 0.93$, and $Median_{task4} = 4$, $SD_{task4} = 1.01$) while Task 1 Lego was rated as half realistic ($Median_{task1} = 4$, $SD_{task1} = 1.10$) and Task 3 Dance as not realistic ($Median_{task3} = 3$, $SD_{task3} = 1.34$). The Kendall's tau correlation test revealed no significant correlations between how realistically people rated a task and their choices of their preferred condition. However, the test indicated a statistically significant positive correlation between Task 3 Dance ratings of how realistic the task is, and Task 4 Biscuits rating ($\tau_b = 0.246, p = 0.039$). So the more realistic people thought Task 3 Dance was, the more they thought Task 4 Biscuits was also. There is also a strong positive correlation between Task 2 Doctor ratings and Task 4 Biscuits ratings of how realistic the task is ($\tau_b = 0.562, p < 0.001$). The most interesting results is that there is a significant positive correlation between how realistically people rated a task, and how critically they rate the same task (see Figure 22). So the more realistic a task is rated, the more critically the task will be rated.

	Task 1 criticality rating	Task 2 criticality rating	Task 3 criticality rating	Task 4 criticality rating
Task 1 realism ratings	0.290* 0.014	0.166 0.169	0.239* 0.042	0.043 0.716
Task 2 realism ratings	0.224 0.064	0.542** 0.000	0.160 0.187	0.426** 0.000
Task 3 realism ratings	0.148 0.199	0.129 0.276	0.495** 0.000	0.164 0.160
Task 4 realism ratings	0.083 0.488	0.437** 0.000	0. 337** 0.005	0.468** 0.000

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Figure 22: Correlation table between ratings of task criticality and task realism.

To conclude we cannot say that the participant's preference of the robot's level of autonomy is independent of the type of task performed by the robot as we found significant associations between the way participants classified tasks and their choice of preferred condition. Therefore hypothesis H4b was verified.

5.6 R5. Experience and knowledge of technology

To measure participants' experience of technology, participants were asked what type of technology they use and to estimate how often they use those everyday technology. Based on participants' answers, we scaled the frequency of usage of technology as such: 0= do not have one, 1= 2- 3 times a month, 2= 2- 3 times a week, 3=more than 3 times a week, 4=less than 30 min a day, 5=1- 2 hours a day, 6=2- 3 hours a day, and 7=More than 3 hours a day. Then we asked participants to rate their familiarity with robots on a scale of 1 to 5 (1 being not familiar at all and 5 being very familiar). As a results, 88% of our participants own a smartphone (98% of them would use it everyday), 86% possess a computer or a laptop they would use everyday, 32% own a tablet they would use everyday, 6% have a smartwatch they use everyday and 12% use a Google Home or Alexa everyday. Regarding experience with robots, the mean was very low regarding familiarity with programming robots ($Median = 1$, $SD = 1.06$), experience with programming robots ($Median = 1$, $SD = 0.91$) and familiarity with interacting with robots ($Median = 1$, $SD = 1.03$). When asked how often participants interacted with robots, half of them

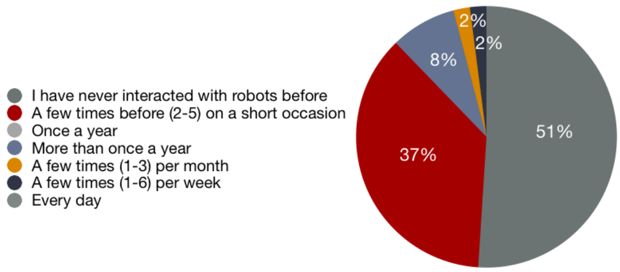

Figure 23: Number of times participants experienced interacting with robots before the experiment.

said never, and 36% of them mentioned on a few occasions before (see Figure 23).

A Kendall's correlation test revealed statistically significant correlations between the frequency of the usage of certain technology and the choice of the preferred conditions participants had for each task (see Figure 24). There is a significant negative correlation between how much time people spend on their computer and people's choice of the preferred condition for Task 1 ($\tau_b = -0.283, p = 0.030$). This means the more often people spend time on their computers, the less they will want to be in control of the robot, when the robot is giving instructions on how to build a Lego character. There is also a significant negative correlation between how often people spend time on their smartphone and people's preferred condition for Task 2 ($\tau_b = -0.306, p = 0.017$). So the more time people spend on their smartphone, the less they want to be in control of the robot when the robot is booking a doctor appointment. The same correlation was found for smartwatch ($\tau_b = -0.368, p = 0.009$). However, the same correlations for Task 3 and Task 4 were not found. Instead, there is a significant positive correlation between how often people use their tablet and their preferred condition for Task 3 ($\tau_b = 0.247, p = 0.049$). So the more people use their tablet, the more they want to be in control of the robot when the robot is dancing.

To conclude, there is a correlation between participants' technology savviness and their preferred level of autonomy of the robot. We found that for both cognitive tasks, the more participants spent time on their smartphone or computer, the less they want to be in control of the robot. However, our data did not indicate that experience with robots influenced participants' preferences for level of autonomy of the robot.

	Frequency usage of smartphone	Frequency usage of computer/ laptop	Frequency usage of tablet	Frequency usage of smartwatch	Frequency usage of Alexa/ Google Home
Task 1 preferred condition	-0.107 0.404	-0.283* 0.030	0.085 0.514	0.138 0.323	-0.089 0.522
Task 2 preferred condition	-0.306* 0.017	0.024 0.856	0.117 0.370	-0.368** 0.009	-0.042 0.761
Task 3 preferred condition	-0.006 0.962	0.011 0.934	0.247* 0.049	0.050 0.707	0.045 0.735
Task 4 preferred condition	-0.077 0.545	-0.043 0.738	-0.248 0.056	-0.059 0.671	-0.085 0.536

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Figure 24: Correlation table between the usage frequency of technology and participants' preferred conditions.

6 Limitations of the study

One of the main difficulties in the design of the experiment was to implement the 4 tasks with the 2 conditions in a consistent way. One of the weaknesses of this study is the Wizard-of-Oz set up. Although the navigation of the robot was fully autonomous, voice and gesture recognition were not implemented in the robot due to time constraints. The selection of physical tasks for the study may also have created some issues. For example, the choice of dancing as a low criticality task to perform for the robot. Many participants were confused into how Sunflower would perform a dance, and some did not see the movements of the robot as dancing movements. In addition, although the high criticality tasks were previously validated [14], we could not use more extreme tasks (such as Sunflower helping to put off a fire in the kitchen) for ethical reasons, as the University Ethics Committee forbids any experiment that may endanger participants physically or/and mentally. As such, it was not possible to use glasses instead of biscuits for Task 4. In extreme cases, participants may have reacted differently [30]. Task 4 was problematic also because some participants were able to guess that the robot was not acting on its own, and also because the task itself became tedious and therefore, it was then more about how well people manoeuved the robot, rather than the mode of the robot (being semi-autonomous or autonomous). The other problem of Task 4 was the expression of happiness used by the robot after carrying biscuits. It was the only condition in which the robot explicitly mentioned happiness and this may have influenced participants. Also, the time spent between each task for participants may have influenced their choices. The task that took the most time on average was Task 1, building a Lego character. A number of participants struggled with the task and a few confessed not having much experience with Lego, or not having good spatial awareness which made the task particularly difficult. Again in this case responses may have been based more on the particular ability of the participant that was tested, rather than the ability of the robot to help in either conditions.

Although the set of questionnaires were mostly standardised, some followed a 5 point Likert Scale and others a 7 point Likert Scale, which might have causes problems regarding comparisons for statistical analysis. In the future, it will be considered to adapt the 5 point Likert Scales to a 7 point to avoid the problem.

7 Discussion

In this study we investigated how much in control of the robot participants want to be depending on the task the robot was performing.

7.1 Preferred condition

People that preferred the autonomous condition C1 mentioned that they felt that Sunflower was more interactive, could do its own dance and demonstrated intelligence. However, some people that chose "other" said that they had no preferences, one said that the C1 autonomous condition was preferable at first due to the fun of the unpredictability of the movements, but would prefer overtime having a robot that does as it is told. Some participants were not able to distinguish conditions and therefore picked the "other" category for this task. This is probably due to the design of the task. The scenario might have been unclear that the aim for the participant was to dance together with the robot. As such people may have expected the robot to always follow their movement as they thought they were teaching the robot dancing, or maybe the limitations of the movements of the robot may have made it difficult for some people to understand the movement of the robot. Task 4, carrying biscuits, clearly showed that participants preferred the autonomous condition C1 when Sunflower decided to carry the biscuits to the living room on its own. Participants mentioned several reasons for it. Some said it was more comfortable to not micromanage the robot and a lot easier than the C2 semi-autonomous condition, or that the robot was faster in this condition. However, some also said that they found it rather difficult to control the robot in the C2 semi-autonomous condition, and some participants even said that the right and left movement seemed to confuse the robot or the robot was not following the instructions. This could be because the C1 autonomous condition was set by the Wizard-of-oz operator, and therefore there was a delay between the voice commands and the robot's movements. However, the robot's navigation was autonomous. Therefore, even if the experimenter clicked on the specific direction where the robot was instructed to go by the participant, the robot may have preferred to take a shorter route. For example, if the participant said to the robot "turn right", if the robot considers that it is more efficient to turn left before turning right, the robot will turn left before turning right, which may give the impression to the participant that the robot is not following their instructions.

7.2 Effect of personality on perception of control

There is no correlation between how controlling people are and how autonomous they wanted the robot to be for any of our tasks, so hypothesis H1 was not verified. However, it was found that the more controlling people are, the less they perceive to be in control of the action in their preferred condition when the robot booked a doctor appointment. It was also found for the same action, that the more controlling people are, the less they felt in control of the robot in their preferred condition. Therefore it can be said that for a high critical cognitive task, controlling people are less likely to feel in control of the robot and the action the robot is performing, even if the robot is acting the way people preferred (being fully autonomous by taking all the decisions or being semi-autonomous, waiting for the user to take the decisions). It could be that since this task (Task 2 booking a doctor appointment) was also rated as the most realistic task, people could relate to it more easily and imagine its consequences better than for other tasks, which would explain why we did not find this correlation for other tasks. Some personality effects were found across every tasks, such as having a negative correlation between Openness and the users' preferred condition for each task. The more openminded people are, the less they are willing to control the robot, which means they are more likely to prefer the robot to be autonomous. These findings complement studies by Meerbeek et al. [31] where they found that the personality of the robot influenced users' preferred levels of control of the robot. However, personality test results have to be taken with caution, as the literature indicates that for previous studies some people apparently preferred robot that express a similar personality [32] while others have shown the opposite [33]. It is suspected that this difference comes about the main task functions the robot expresses. As this live experiment demonstrated, depending on the type of task the robot performs, users have different expectations from the robot.

7.3 Perception of control, task criticality and type of task

7.3.1 Low critical tasks

The results show that for Task 1, building a Lego character, and Task 3, doing a dance, participants preferred the C2 semi-autonomous condition, when the robot was following their instructions. It could be because participants felt that the success of Task 1 was depending more on their skills as they instructed the robot to show the next task. Therefore, it could then be that participants considered that it was more important to do the task correctly when they felt in control of the robot. This contradicts some early research done by Meerbeek et al.[31] where they found that when the robot was performing as a TV assistant task (choosing a TV channel), his participants showed no preference for the user's level of control. But as the researchers mentioned in their article, they focused on the contrasting personality behaviours that the robot was displaying, therefore this has to be taken with caution. Also, it does not appear that the participants were asked to rate the criticality of their task in this study. The results of this experiment regarding the low criticality tasks are consistent with the findings from our first live experiment [12], where the robot companion was performing a cleaning task, which was considered, according to our questionnaire study [14] as a task low in criticality. The more controlling participants are, the more autonomously they wanted the robot to be. The results of this live experiment are not only consistent with the previous results but also reinforced by the fact that 50 people took part in the study while only 33 people took part on the Meerbeek's study that was conducted more than 10 years ago. Therefore it seems that for tasks low in criticality, participants prefer to be in control of the robot.

7.3.2 High critical tasks

As our results show for Task 2 Doctor and Task 4 Biscuits, the more the participants felt in control when the robot was in charge (autonomous condition C1), the more the participant was willing to let the robot be in charge. Perhaps this was the case because Task 2 Doctor and Task 4 Biscuits, are both considered critical tasks and therefore, it is more of a relief for the participant if the robot can actually do the task correctly. Also, it could be that people felt more in control of the action as the robot was navigating faster and smoother than in the other condition, when the participant had to guide the robot. As Task 4 Biscuits is a physical task, Sunflower had to move whenever a command was said and it could be that the distance was not matching what people expected or simply as one participant told the experimenter, they could see that the robot was not responding to its sensors, therefore knew in this particular set up that the robot was remote-controlled.

However, the data also shows that despite this result, the majority of our participants preferred the C2 semi-autonomous condition for Task 2 Doctor while they preferred the C1 fully autonomous condition for Task 4 Biscuits. Task 2 Doctor is a cognitive task which is about scheduling participants day to day life. Therefore, as many mentioned, they "need to have control of the situation", they "want to make [their] own decisions based on the options available" and they "like to make the final choice". When Sunflower booked the appointment for the participant in the C1 fully autonomous condition, the robot picked the first available slot on the digital calendar. This means that if for a particular reason, this day the participant wanted to leave the day open for other plans, they would have had to mention it in the digital calendar Sunflower was referring to. This means that it would have been more hard work for the participant to implement these plans and ideas in the calendar, and to let Sunflower know rather than for the participant to take an overall decision based on an overview of potential events/appointments that could happen that day. Some participants felt they were "infantilised", which is one of the risks mentioned by early research on ethics [5]. This research confirms to some extent the danger mentioned by Lucidi [34]. This is most probably why almost every participants did not prefer the robot to take the decision for the doctor appointment as this relates to how people manage their life.

However, the same results did not apply for the high critical physical task "carrying a biscuit". For Task 4, carrying biscuits, people preferred the autonomous condition C1 as they mention it was difficult to manoeuver the robot with voice commands in the C2 semi-autonomous condition but also because many said it was easier this way ("it is more comfortable to not micromanage it") and that the robot demonstrated more intelligence. Task 4 Biscuits being a physical task, people could also see the robot carrying the biscuits from the kitchen to the living room and therefore had an immediate overview of what is happening. So even if the robot is not as accurate as they would like it to be regarding the exact location of where the biscuits should be carried in the living room, the main task of carrying the biscuits was accomplished. When it comes to a physical task, people do not expect to have to supervise and micromanage it, at least in this case with carrying biscuits, because it is more hard work for them to do the supervision rather than to do the task itself. For the cognitive task, it is the opposite. Therefore it is most probably why participants preferred the C1 fully autonomous condition as they felt in control of the situation.

Those results could also be explained by how realistically people perceived the task to be. As Task 2 Doctor was perceived as more realistic than Task 4 Biscuits, (although both were rated as very realistic on average). Because people related more easily to Task 2 Doctor, this could also be why participants prefer the C2 semi-autonomous condition.

So our results here have demonstrated that people would be willing to let the robot being fully autonomous if they feel in control of the situation. This is obviously more difficult to implement for cognitive tasks, as it is harder to adapt to each person's habits.

7.4 Effect of technology knowledge and awareness

The results revealed that the more time people spend with their everyday technology, the more autonomous they want the robot to be. This could be the case because Sunflower's interface was a tablet. Since the tablet is a widely used and available technology, people are used to manipulate one, or at least something that is similar, such as a smartphone or a computer. If the robot's interface would have been something less familiar, people might have preferred the robot to be less autonomous. So our results confirm there is an habituation effect which can be exploited by roboticists. However, this is only true for smartphones and computers. There was no such findings for people having previous experience with robots. It could be because the majority of our participants were inexperienced robot users, therefore it is difficult to draw conclusions with such data.

8 Conclusion

To conclude, our results show there is a strong correlation between people's perception of control and their choice of how autonomous the robot should be. However, our results did not exactly corroborate our findings from our previous investigations [12]. As suspected, people's preferences for level of robot autonomy are subtle. The type of task and the criticality of the task influences the way people want their robot to react. For a critical cognitive task, people prefer the robot to be semi-autonomous, so they could take the final decision. Whereas, for a critical physical task, people prefer the robot to be autonomous as they did not want to micromanage the robots movements while it was performing the task. The results demonstrated that people would be willing to let the robot being fully autonomous provided they feel in control of the situation. This is more difficult to implement for cognitive tasks than for physical tasks, as it is harder to predict people's intentions and habits. It would be good to investigate in the future if we can change people's preferences for cognitive critical tasks, by having a robot that would be more inquisitive. Also, it is necessary to conduct experiments with more realistic physical tasks that the robot would perform for participants, as our selection of tasks was revealed to be too ambiguous for participants to really relate to.

References

[1] IRobot, https://www.irobot.com/roomba (Last accessed 25/06/2019)

[2] J.-Y. Sung, L. Guo, R. E. Grinter, H. I. Christensen, "My Roomba Is Rambo": Intimate Home Appliances, Springer, 2007

[3] Moulinex, https://www.moulinex.fr/Cuisson/Robotcuiseur-Companion/c/cooking+food+processors (Last accessed 29/01/2019)

[4] Ubtech, https://www.theverge.com/2017/11/20/16681396/amazon-alexa-powered-lynx-robot-ubtechrobotics (Last accessed 29/01/2019)

[5] B. Whitby, Sometimes it's hard to be a robot: A call for action on the ethics of abusing artificial agents, Interacting with Computers, 2008, 20(3), 326–333

[6] B. Whitby, Do you want a robot lover? the ethics of caring technologies, Robot ethics: The ethical and social implications of robotics, 2011, 233–248

[7] A. Sharkey, N. Sharkey, Granny and the robots: ethical issues in robot care for the elderly, Ethics and information technology, 2012, 14(1), 27–40

[8] J. Bernotat, F. Eyssel, Can ('t) wait to have a robot at home?-japanese and german users' attitudes toward service robots in smart homes, In: 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), IEEE, 2018, 15–22

[9] B. Shneiderman, P. Maes, Direct manipulation vs. interface agents, Interactions, 1997, 4(6), 42–61

[10] R. Pekelney, R. Chu, Design criteria of an ergonomic mouse computer input device, In: Proceedings of the Human Factors and Er-

gonomics Society Annual Meeting, SAGE Publications, Sage CA: Los Angeles, CA, 1995, 39(5), 369–373
[11] M. W. Gallagher, K. H. Bentley, D. H. Barlow, Perceived control and vulnerability to anxiety disorders: A metaanalytic review, Cognitive Therapy and Research, 2014, 38(6), 571–584
[12] A. Chanseau, K. Dautenhahn, K. L. Koay, M. Salem, Who is in charge? sense of control and robot anxiety in human-robot interaction, In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), IEEE, 2016, 743–748
[13] P. Haggard, V. Chambon, Sense of agency, Current Biology, 2012, 22(10), R390–R392
[14] A. Chanseau, K. Dautenhahn, M. L. Walters, K. L. Koay, G. Lakatos, M. Salem, Does the appearance of a robot influence people's perception of task criticality?, In: 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), IEEE, 2018, 1057–1062
[15] Google Duplex, https://www.androidauthority.com/ultimate-it-certification-bundle-983274/ (Last accessed 25/06/2019)
[16] E. Pacherie, The sense of control and the sense of agency, Psyche, 2007, 13(1), 1–30
[17] H. M. Lefcourt, Locus of control, Academic Press, 1991
[18] H. A. Yanco, J. L. Drury, A taxonomy for human-robot interaction, In: Proceedings of the AAAI Fall Symposium on Human-Robot Interaction, 2002, 111–119
[19] S. G. Tzafestas, Human-robot social interaction, In: Sociorobot World, Springer, 2016, 4, 53–69
[20] J. Guiochet, M. Machin, H. Waeselynck, Safety-critical advanced robots: A survey, Robotics and Autonomous Systems, Elsevier, 2017, 94, 43–52
[21] J. Dai, P. Taylor, H. Liu, H. Lin, Folding algorithms and mechanisms synthesis for robotic ironing, International Journal of Clothing Science and Technology, 2004, 16(1/2), 204–214
[22] N. Ezer, A. D. Fisk, W. A. Rogers, More than a servant: Self- reported willingness of younger and older adults to having a robot perform interactive and critical tasks in the home, In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications, Sage CA: Los Angeles, CA, 2009, 53(2), 136–140
[23] S. D. Gosling, P. J. Rentfrow, W. B. Swann Jr., A very brief measure of the big-five personality domains, Journal of Research in personality, 2003, 37(6), 504–528
[24] L. E. McCutcheon, The desirability of control scale: Still reliable and valid twenty years later, Current research in social psychology, 2000, 5(15), 225–235
[25] Lego questionnaire, https://docs.google.com/forms/d/1FJc4ZE4teRYrGjLzih39MYWvQVZGEFLkwoU4tmiBNzI/viewform, (Last accessed 25/06/2019)
[26] Doctor questionnaire, https://bit.ly/2Xz70d2 (Last accessed 25/06/2019)
[27] Dance questionnaire, https://docs.google.com/forms/d/e/1FAIpQLSc9vHeoRVs9y6vlDDJyH0iA_hPKk_fQdkjOLIeEw9SUVV8_KQ/viewform (Last accessed 25/06/2019)
[28] Biscuits questionnaire, https://docs.google.com/forms/d/e/1FAIpQLSeT6SqeRUIFun0mIeuvYnmLEvfIPbXQt4IF5dHaYDVUPBPEw/viewform (Last accessed 25/06/2019)
[29] J. M. Burger, H. M. Cooper, The desirability of control, Motivation and emotion, 1979, 3(4), 381–393
[30] P. Robinette, A. M. Howard, A. R. Wagner, Effect of robot performance on human-robot trust in time-critical situations, IEEE Transactions on Human-Machine Systems, 2017, 47(4), 425–436
[31] B. Meerbeek, J. Hoonhout, P. Bingley, J. M. Terken, The influence of robotA. Tapus, C. Ţăpuş, M. J. Matarić, User-robot personality matching and assistive robot behavior adaptation personality on perceived and preferred level of user control, Interaction Studies, 2008, 9(2), 204–229
[32] for post-stroke rehabilitation therapy, Intelligent Service Robotics, 2008, 1(2), 169–183
[33] S. Woods, K. Dautenhahn, C. Kaouri, R. te Boekhorst, K. L. Koay, M. L. Walters, Are robots like people?: Relationships between participant and robot personality traits in human-robot interaction studies, Interaction Studies, 2007, 8(2), 281–305
[34] P. Bisconti Lucidi, D. Nardi, Companion robots: the hallucinatory danger of human-robot interactions, In: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, ACM, 2018, 17–22