

Does the appearance of a robot influence people's perception of task criticality?

Adeline Chanseau, Kerstin Dautenhahn, Michael L. Walters, Kheng Lee Koay,
Gabriella Lakatos and Maha Salem
Adaptive Systems Research Group, School of Computer Science
University of Hertfordshire

Abstract—As home robot companions become more common, it is important to understand what types of tasks are considered critical to perform correctly. This paper provides working definitions of task criticality, physical and cognitive tasks with respect to robot task performance. Our research also suggests that although people's perceptions of task criticality is independent of robot appearances, their expectation that a robot performs tasks correctly is affected by its appearance.

I. INTRODUCTION

With the current popularisation of devices such as Google Home and Amazon Alexa, it is important to distinguish what people consider critical for a domestic robot companion to do reliably. Previous research ([1], [2]) suggests that evaluating the criticality of a task is difficult because of the lack of standardisation in the field. To tackle this problem, this study investigated how people defined task criticality, and whether there is a relationship between the (subjective) level of criticality people attributed to a task and the appearance of the robot performing the task. This work proposes a definition of task criticality, and also a definition of cognitive and physical tasks performed by robots. We also investigated if people's perception of task criticality depends on the appearance of the robot carrying out the task.

II. BACKGROUND AND MOTIVATION

A. What is criticality?

Criticality is an unclear concept that has been widely studied in different areas of research. In linguistics, criticality is defined as "an evaluative judgement made within any field of human activity about some aspect, object or behaviour of that field" [3], meaning that criticality is subjective and context dependent. In biology, criticality "describes sudden changes in the state of a system when underlying processes change slightly" [4]. Criticality is then perceived as a sudden dramatic change in the expected event [4], [5]. In the field of Human-Robot Interaction (HRI), Yanco and Drury [6] defined criticality as "the importance of getting the task done correctly in terms of its negative effects should problems occur". Since this definition was provided, no research has yet been performed to analyse how to apply it to standard tasks performed by robots, and more precisely, everyday tasks that home robot companions might be expected to perform. The current study investigated people's own definition of task criticality for domestic robot companions. Firstly, to validate and update the definition given by [6], and secondly, to understand better what influences people's perceptions of task criticality. In order to apply the definition of task criticality

in practice, it is important to be able to distinguish between different levels of criticality. Tzafestas [1] described three levels of criticality in his research: low, medium and high, without providing any guidelines into how to distinguish these levels. The study described in this paper aims to provide guidelines as to what makes a task more or less critical. Some previous studies [7]–[9] have suggested that the level of perceived task criticality performed by a robot depends on how much people wanted the task to be performed by a human. Beer et al. [7], [8] and Mitzer et al. [9] used the example of giving medication in their studies and found that people preferred having a human for this task rather than a robot. However, their findings suggested that some other factors may have influenced their participants, such as trust or the visual appearance of the robot. Therefore our study also investigated if there is a relationship between how participants rate task criticality and robot appearance.

B. How does the appearance of the robot affects people's perception?

Several previous studies have shown that robot appearance affects the people's judgement of robot behaviour. In one of the early studies, Goetz et al. [10] investigated how to improve Human-Robot Cooperation by matching robot appearance and behaviour to the task the robot had to complete. Later on, Walters et al. [11] showed that there is a tendency for people to prefer some human-like attributes in robots. In a recent study, Malle et al. [12] demonstrated that robot appearance can also affect people's moral judgements about robots. In a moral dilemma, people blame robots more for inaction than action, and they blame humans more for the opposite. They also found evidence that people treated a mechanical-looking robot differently from a human-looking one, when both robots were described identically. Abubshait and Wiese [13] investigated how robot appearance and behaviour influence HRI and they found that a robot's appearance affects mind judgements (e.g the attributed intentions the robot has). Salem et al. [14] suggested in their study that "the robot's level of anthropomorphism may lead to different degrees of 'forgiveness' in humans". Although their study did not focus on appearance, as only one robot was used, potentially this could mean that the more human-like the robot appears, the less forgiving people will be when it makes mistakes. This can be linked to findings by Mitzer et al. [9] where the level of task criticality depends on how preferable it is for the user

to have a human performing a critical task, such as giving medication.

C. Perception of tasks performed by the robots

Previous research has investigated how trustworthy a robot is perceived by people, depending on the task the robot is performing. Salem et al. [14] showed in their study that the type of task performed by the robot matters. It seems that the irreversibility of some actions to carry out by people which were suggested by the robot prevented most of their participants from performing them. Prakash and Rogers [15] showed that perceptions of robots' human-likenesses changes for different types of tasks (personal care, social, decision-making and chores). Their experiment underlined that robotic appearance was least appreciated for decision-making types of tasks, in their case money investment. Overall they found that older people preferred human-like robots for personal care, chores, social and decision-making tasks compared to younger people who expressed a more diverse preference (mechanical appearance, mixed appearance or human appearance). Hinds et al. [16] found that in an industrial context where robots and humans work together on various tasks (assembling objects, carrying objects, designing something with the participant), people preferred overall to have a machine-like robot over a human-like robot. This difference shows that robot appearance preferences may depend on the environmental context (e.g. home versus a factory). The current study investigated whether task criticality depends on the robot's appearance.

III. METHOD

To investigate how people rate task criticality according to the robot appearance, a questionnaire-based study was conducted.

A. Research questions

- R1. What defines a cognitive task versus a physical task for a domestic robot companion?
- R2. What defines task criticality?
- R3. Are people's perceptions of task criticality influenced by the robot's appearance?

B. Experimental procedure

Four different questionnaires were prepared (see [17]), each containing identical questions, and each showing a different picture of a robot companion, for participants to imagine what the robot looks like when performing a given list of tasks. Each participant received one of the four questionnaires randomly. The questionnaire had five sections: demographics, usage of technology, people's expectations of a robot companion, rating of task criticality and defining task criticality. In a previous pilot study, when participants responded to an open-ended question asking to define criticality, many expressed confusion and difficulty in expressing the concept. Therefore, in the current study, we deliberately chose not to provide a definition of criticality in the study beforehand in order not to bias the participants, and instead to get participants' own definitions of criticality via a small

set of statements that they had to rate on 5-point Likert scales (see Table IV). Open-ended questions were used for specific tasks to provide better context and were then classified according to keywords in the analysis.

Participants were recruited from University staff and students and through social media. As a result, 84 people completed the questionnaire (35 female and 49 male). Their ages ranged between 19 and 64 ($M = 35$, $SD = 12.221$). There were 22 people that answered the questionnaire showing the Sunflower robot picture (Fig.1a), 21 people showing the Pepper robot picture (Fig.1b), 21 people showing the Sawyer robot picture (Fig.1c) and 20 people showing the combined Sunflower and Roomba picture (Fig.1d). Participants were asked to rate the robot's appearance on a scale of 1 to 7 (1 being very-machine like and 7 being very human-like), apart from the ones who had both Roomba and Sunflower as robots pictures in their questionnaire. It was further decided to provide in one of the questionnaires a picture of both Sunflower and Roomba to see if the perception of the robot companion (Sunflower) changed with the presence of another robot that was task orientated.

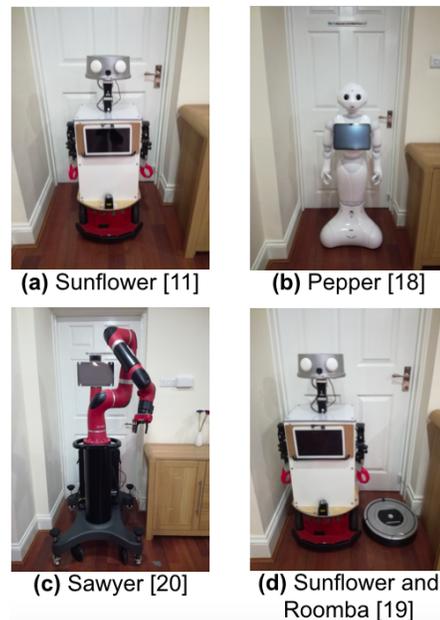


Fig. 1: Robot pictures shown to the participants

IV. RESULTS

A. R1. Definition of cognitive and physical tasks

Participants were asked to define what they considered physical and cognitive tasks. Their definitions were classified according to recurrent keywords mentioned by the participants. For the definition of physical tasks, Table I, people mentioned, regardless of their questionnaire body, movements, strength and objects. It can be noted that none of the participants who were shown the Sawyer robot mentioned anything related to the body, while the majority of the participants who had Sunflower and Roomba as a picture mentioned the necessity of force in their definition of physical tasks. It also shows that participants that had Sunflower

and Roomba as a picture focussed more on the Roomba robot for their definition, than the ones that only had the Sunflower picture. As a result, we can define for robots, a physical task as any task that requires body movements or motion, which may be qualified as a laborious task.

TABLE I: Definition of physical tasks depending on the image of the robot provided

Key words	Number of participants mentioning these key words			
	Sunflower	Pepper	Sawyer	Sunflower + Roomba
body (requires a body/body parts, embodiment, artificial/natural body...)	4	5	0	3
movement (requires to move, motion involved ...)	8	6	1	5
strength (requires force/effort, involves manual tasks...)	3	7	6	7
interaction with objects and or the environment	4	4	3	1

For the definition of a cognitive task, participants mainly mentioned a mind process, information analysis, decisions or qualify it as an antonym to physical task (see Table II). For robots, a cognitive task can therefore be defined as any task that requires mental activities or thinking processes and which may involve some decision making. Participants who were shown a picture of both Sunflower and Roomba, mainly mentioned information processing in their definition of a cognitive task for a robot, which again shows that these participants were more focussed on the Roomba robot than the ones who had the Sunflower only questionnaire. It might be that Roomba being a commercially available robot, participants may have more familiarity with it. Also, the Roomba being mainly a physical robot (vacuum cleaning being its sole purpose), people considered some cognitive aspects of cleaning such as "being able to distinguish a carpet from a tiled floor".

TABLE II: Definition of cognitive tasks depending on the image of the robot provided

Key words	Number of participants mentioning these key words			
	Sunflower	Pepper	Sawyer	Sunflower + Roomba
thinking (involves mental process, mind/thoughts ...)	9	8	5	3
information processing (requires analysis ...)	6	2	3	11
making use of the brain (decision making ...)	2	3	3	3
non-physical interaction	1	1	4	1

These definitions are supported by the way participants classify as either "physical", "cognitive", "both" or "other please specify", a list of tasks (see Table III) the robot could

do for them. Tasks that clearly involve motion were classified either as physical or both (A. vacuuming, C. dancing, F. carrying or K. waving) and tasks that involve thinking as cognitive.

B. R2. Definition of task criticality

The Pearson Chi-square test showed there is no relationship between the classification (cognitive or physical) of a task and the level of its criticality ($\chi^2 = 0.400$, $df = 2$, $p < 0.819$). Participants were asked why they chose to classify a task as highly critical or not highly critical. Most of the reasons why participants rated tasks as highly critical were related to some dimension of risk. For example the potential harm of another person, impact related to health or money related ("expensive [champagne] flutes at risk"), and potential social impact such as "punctuality for interview". Tasks that were rated low critical were those that had low impact with reversible consequences, such as vacuuming or tasks focussing on entertainment. This is consistent with the results presented in Table V, illustrating factors which are taken into consideration for criticality. To investigate factors that people consider when evaluating the criticality of a task, they were presented a list of statements (see Table IV) and asked to rate on a scale of 1 to 5 (1 being not important for criticality at all and 5 being very important for criticality), which aspects they considered important to judge for the criticality of a task.

The results show people considered mainly four aspects when judging task criticality: the task being carried out safely, the importance of the task, the task being carried out correctly and the task being carried out with attention to detail. Therefore task criticality can be defined as the importance of a task being carried out safely, correctly and with attention to detail. Due to the low sample size for each set of questionnaire (each fewer or equal to 22 participants), and the lack of balance between gender and age, we could not apply the test for normal distribution. It was therefore chosen to perform non-parametric tests. A non-parametric Kendall's tau correlation test showed that there is a significant positive correlation ($n = 84$, $\tau = 0.246$, $p < 0.008$) between how people rated a task being carried out in a timely manner and the task being carried out with attention to detail. So the more important it is that a task has to be carried out in a timely manner, the more important it is that it is done with attention to detail. Figure 2 shows participants were consistent with their answers. There is significant positive correlation ($n = 84$, $\tau = 0.325$, $p < 0.001$) between how people rated a task being carried out with attention to detail and the difficulty of the task.

This result is also consistent with how people prioritised types of tasks (see Table V). Participants were asked to rank the statements in Table V from the most important thing the robot can do (rank 5) to the least important one (rank 1). As a result, "security" was consistently rated as the most important task and "entertainment" as the least important one (see Table.VI), when participants were asked "which

TABLE III: Classification of the type and the criticality of the tasks

Classification of task type and its percentage rating		Tasks	Classification of task criticality and its percentage rating		Chi square test with robot appearance
Physical	50%	M. Your paper bin is full. The robot is taking out the trash for you.	Low	61.9%	$\chi^2(10) = 11.965$ $p = 0.287$
	54.8%	A. There is some confetti on your living room floor. The robot is vacuuming confetti off the floor.	Low	71%	$\chi^2(10) = 16.895$ $p = 0.077$
	54.8%	C. You are sitting on the sofa, relaxing. You want to see a dance performance. The robot is performing a dance to entertain you.	Low	72.6%	$\chi^2(10) = 0.611$ $p = 0.611$
	61.9%	F. You want to transport some fragile crystal champagne flutes to the living room. The robot is transporting the glasses you cannot carry.	High	59.5%	$\chi^2(10) = 19.928$ $p = 0.030$
	67.9%	O. You have some hungry guests in the living room. The robot is helping you carrying appetizers from the kitchen to the living room.	High	45.2%	$\chi^2(10) = 13.363$ $p = 0.204$
Cognitive	65.5%	D. You need to prepare a drink for your sister's six-month-old baby. The robot is reading to you the instructions of the recipe sent by the mother.	High	58.3%	$\chi^2(10) = 6.400$ $p = 0.781$
	76.2%	P. You want to cook a new recipe sent by your friend for dinner. The robot is reading to you the instructions of the recipe.	Low	59.5%	$\chi^2(10) = 22.044$ $p = 0.015$
	83.3%	H. Your interview is upcoming. The robot is reminding you of the name of the company and the person you will meet with a short description of their profiles the day before the interview.	High	53.6%	$\chi^2(10) = 8.965$ $p = 0.535$
	84.5%	B. You have just remembered that you need to see the doctor this week for a blood test. The robot is helping you by checking your availability on your diary and booking a suitable appointment with the doctor via the Internet.	High	73.8%	$\chi^2(10) = 14.133$ $p = 0.167$
	85.7%	J. You want to send some flowers to your partner for Valentine's day. The robot is helping you ordering flowers online by showing a selection of your partner's favourite flowers and what time the selected bouquet is guaranteed to be delivered at.	Low	45.2%	$\chi^2(10) = 14.693$ $p = 0.144$
	85.7%	L. Nobody has watered your plants today. The robot is reminding you to water the plants by sending you a notification.	Low	56.0%	$\chi^2(10) = 11.965$ $p = 0.287$
	85.7%	Q. Your job interview is later on today. The robot is calculating the travel time and the best route required to get to the interview and will notify you when it is time to leave to arrive on time.	High	72.6%	$\chi^2(10) = 9.706$ $p = 0.467$
	86.9%	N. You have just received a challenge from your best friend, solving a deconstructed 3D wooden puzzle in less than 5 minutes. The robot is offering to help you to solve the puzzle by giving you clues and showing you pictures of the constructed puzzle.	Low	63.1%	$\chi^2(10) = 6.326$ $p = 0.787$
Both Physical and Cognitive	46.4%	R. You have set up your alarm clock to wake up in the morning to catch a flight. You give to the robot your alarm clock so the robot can move the ringing alarm clock in the morning to force you out of bed to stop the alarm clock.	High	59.5%	$\chi^2(10) = 4.553$ $p = 0.919$
	50%	K. Some visitors have arrived. Your robot approaches them and greets them cheerfully by moving in a circular motion.	Low	57.1%	$\chi^2(10) = 11.886$ $p = 0.293$
	57.1%	I. There is a mess in the living room, your six-year-old nephew left his toys everywhere. The robot helps you collecting the toys and putting them into a box.	Low	63.1%	$\chi^2(10) = 13.073$ $p = 0.220$
	66.7%	E. You have lost your car keys and need to drop off your friend at the train station immediately. The robot is looking for your car keys by moving around the apartment and scanning the area.	High	67.9%	$\chi^2(10) = 8.186$ $p = 0.611$
	75%	G. Your hamster pet escaped from its cage and got lost in the house. The robot is helping you looking for the pet by moving around the house, and scanning different rooms.	High	66.67%	$\chi^2(10) = 11.740$ $p = 0.303$

aspects do you consider important for judging the criticality of a task?" Although "security tasks" were rated as the most critical type of tasks across all sets of questionnaires, there was a noticeable difference depending on the image of the robot participants had viewed. Fifty-seven percent of the participants ranked security as the most important factor for Sawyer, which can be explained by the bulkier appearance of the robot, compared to only 41% of the participants who rated security as the most important factor for Sunflower.

C. R3. Perception of tasks influenced by robot appearance

The Kendall's tau correlation test showed a significant positive correlation ($n = 22$, $\tau = 0.396$, $p < 0.05$)

between the Sunflower robot's human-likeness rating and how important it is for the robot to perform a given task correctly. However there was no such correlation for the Pepper robot ($n = 21$, $\tau = -0.44$, $p = 0.814$) and the Sawyer robot ($n = 21$, $\tau = -0.202$, $p = 0.306$). This suggests, the more human-like the Sunflower robot was perceived by participants, the more important they considered that the task performed by the robot should be carried out correctly. Perhaps participants were more likely to consider the importance of the task being carried out correctly, because of the custom-made appearance of the Sunflower's robot, as compared to both Pepper and Sawyer (both manufactured

TABLE IV: How participants scored statements defining task criticality on average

Statements	Sunflower	Pepper	Sawyer	Sunflower + Roomba	Total average
Task being carried out correctly (task being carried out wrongly can lead to irreversible effects such as glass being broken)	4.23	4.10	4.33	4.15	4.20
Task being carried out in a timely manner (task not being carried out in a timely manner could lead to nuisance such as hoovering being done in the living room while you are watching TV)	3.45	3.57	3.29	2.90	3.30
Task being carried out with attention to detail (for example ironing clothes at the right temperature)	3.95	3.90	4.38	3.95	4.05
Difficulty of the task (for example cooking which involves chopping vegetables, heating up a pot of water, etc...)	3.09	3.29	4	3.65	3.51
Importance of the task (for example reminding you to pick up your daughter from school)	4.5	4.52	4.29	4.30	4.40
How personal the task is (for example giving fashion advice)	2.63	2.42	2.14	2.90	2.51
Task being carried out safely in order not to break/damage objects or injure people (e.g. carrying glasses slowly)	4.5	4.62	4.67	4.60	4.60

Fig. 2: Correlations between statements related to participants' criticality rating

	Task being carried out in a timely manner	Task being carried out with attention to detail	Difficulty of the task	Importance of the task	How personal the task is	Task being carried out safely
Task being carried out correctly	0.418**	0.386**	0.024	0.319**	0.002	0.449**
Task being carried out in a timely manner		0.246**	0.097	0.185*	0.120	0.170
Task being carried out with attention to detail			0.325**	0.164	0.221*	0.233*
Difficulty of the task				-0.052	0.150	0.039
Importance of the task					0.067	0.225*
How personal the task is						0.010

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

TABLE V: List of types of tasks for considering what is the most important thing a house robot companion can do

Type of tasks
A. basic household chores (cleaning, taking out trash, vacuuming ...)
B. monitoring the house (checking if the oven is still on, if there is some milk left ...)
C. secretary tasks (acting as a reminder for appointments, setting up appointments, taking messages ...)
D. security tasks (acting as a bodyguard, calling the police when someone tries to break in the house ...)
E. entertainment tasks (displaying a dance to the owner, telling a joke, showing videos...)

TABLE VI: How people prioritised the type of tasks for a robot companion to do

Ranking from the most important to the least	Sunflower		Pepper		Sawyer		Sunflower + Roomba		Total	
	Type of tasks	Number of participants	Type of tasks	Number of participants	Type of tasks	Number of participants	Type of tasks	Number of participants	Type of tasks	Number of participants
Highest priority	D	41%	D	45%	D	57%	D	56%	D	49%
2nd highest priority	B	45%	C	40%	B	47%	B	50%	B	43%
3rd highest priority	C	50%	B	50%	B	33%	C	33%	C	36%
4th highest priority	A	41%	A	40%	AC	43%	AC	39%	A	41%
Lowest priority	E	82%	E	85%	E	90%	E	67%	E	81%

and commercially designed robots). When participants were asked to rate the robot appearance, on a scale of 1 to 7 (1 being very machine-like and 7 being very human-like), Sawyer was clearly classified as machine-like ($n = 21$, $M_{Sawyer} = 2.1$, $SD_{Sawyer} = 1.55$), while Pepper was clearly classified as more human-like ($n = 21$, $M_{Pepper} = 3.76$, $SD_{Pepper} = 1.26$). Sunflower on the other hand was classified between Sawyer and Pepper in terms of human-likeness

appearance ($n = 22$, $M_{Sunflower} = 2.57$, $SD_{Sunflower} = 1.12$). A Pearson's Chi square test showed no association between the way participants rated task criticality and the robot's appearance (see Table IV), with the exception of 2 tasks, carrying champagne flutes and reading a recipe sent by a friend. Perhaps there is an association for the carrying task because of the shape of the gripper/hand of the robot. Similarly for the task of reading a recipe sent by

a friend, the association could be due to the human-likeness of the robot. But the Chi square test showed no significant results for carrying appetizers and reading a drink recipe. There was no statistically significant correlation between ratings of the robot's human-likeness and participants' ages ($n = 64$, $\tau = -0.054$, $p = 0.561$). This means there is no evidence that younger people tend to perceive robots as more human-like. Similarly the correlation tests showed no statistically significant correlations between the amount of time participants had previously spent interacting with robots, or their familiarity with robots, and how they rated the appearance of the robot. However, there was a significant positive correlation ($n = 64$, $\tau = 0.295$, $p = 0.004$) between how human-like participants rated the appearance of the robot, and how much time they had spent programming robots they previously experienced. So it seems that the more time people have spent on programming robots, the more human-like they tended to rate the robot's appearance. This result has to be taken with caution, because of the small number of participants per set (less than 22), and a tendency for the participants to have little programming experience.

V. DISCUSSION AND CONCLUSION

The main outcomes of this study were to show that the perception of task criticality is independent of the robot's appearance, to clarify the definitions of physical and cognitive tasks for a robot, and to define task criticality. As a result, a physical task was defined as **"any task that requires body movement or motion processes, which may be qualified as a laborious task"**. A cognitive task was defined as **"any task that requires mental activities or thinking processes, which may involve some decision making"**. Task criticality can be defined as **"the importance of a task being carried out safely, correctly and with attention to detail"**. The consistency of our findings for criticality shows there is a definite contribution to the community by clarifying how task criticality is perceived for a home robot companion. Moreover, the research highlights the main factors which are considered when assessing for high task criticality (i.e security and safety). For example, entertainment scored low on risks to security and safety, so it was classified as a low critical task.

This paper also showed that the majority of a tasks criticality classification was independent from people's rating of a robot's appearance. If a robot's appearance has the right balance between machine-likeness and human-likeness, the user will tend to focus more on how the robot should perform the task correctly. However further investigations are needed to confirm these results, and it is difficult to evaluate what the right balance is since this finding applied to a custom-made robot, Sunflower. Also the questionnaire study only showed images of the robots. Therefore people did not have an appreciation of how the robots acted dynamically in the real world, which is a limitation of this study. The low number of participants per set of questionnaires for this study is another limitation. But the findings indicate there is definitely

a need to further investigate task criticality with live robots in order to further consolidate and refine the definition of task criticality. It is planned to conduct an experimental study with live robots in the future to confirm these results and to investigate further if task criticality can be linked to factors such as sense of control or trust.

In conclusion, this paper has provided working definitions of task criticality, physical and cognitive tasks, and indicated rating of task criticality is independent of robot appearance.

REFERENCES

- [1] S. G. Tzafestas, "Human-robot social interaction," in *Sociorobot World*. Springer, 2016, ch. 4, pp. 53–69.
- [2] T. Salter, F. Michaud, and H. Larouche, "How wild is wild? a taxonomy to characterize the 'wildness' of child-robot interaction," *International Journal of Social Robotics*, vol. 2, no. 4, pp. 405–415, 2010.
- [3] I. Bruce, "Expressing criticality in the literature review in research article introductions in applied linguistics and psychology," *English for Specific Purposes*, vol. 36, pp. 85–96, 2014.
- [4] M. Pascual and F. Guichard, "Criticality and disturbance in spatial ecological systems," *Trends in ecology & evolution*, vol. 20, no. 2, pp. 88–95, 2005.
- [5] W. D. Bygrave and C. W. Hofer, "Theorizing about entrepreneurship," *Entrepreneurship theory and Practice*, vol. 16, no. 2, pp. 13–22, 1992.
- [6] H. A. Yanco and J. L. Drury, "A taxonomy for human-robot interaction," in *Proceedings of the AAAI Fall Symposium on Human-Robot Interaction*, 2002, pp. 111–119.
- [7] J. M. Beer, A. D. Fisk, and W. A. Rogers, "Toward a psychological framework for levels of robot autonomy in human-robot interaction," Georgia Institute of Technology, Tech. Rep., 2012.
- [8] J. Beer, A. D. Fisk, and W. A. Rogers, "Toward a framework for levels of robot autonomy in human-robot interaction," *Journal of Human-Robot Interaction*, vol. 3, no. 2, p. 74, 2014.
- [9] T. L. Mitzner, C.-A. Smarr, J. M. Beer, T. L. Chen, J. M. Springman, A. Prakash, C. C. Kemp, and W. A. Rogers, "Older adults' acceptance of assistive robots for the home," Technical Report HFA-TR-1105 Atlanta, GA: Georgia Institute of Technology School of Psychology–Human Factors and Aging Laboratory, Tech. Rep., 2011.
- [10] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behavior to tasks to improve human-robot cooperation," in *Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003. The 12th IEEE International Workshop on*. IEEE, 2003, pp. 55–60.
- [11] M. L. Walters, D. S. Syrdal, K. Dautenhahn, R. Te Boekhorst, and K. L. Koay, "Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion," *Autonomous Robots*, vol. 24, no. 2, pp. 159–178, 2008.
- [12] B. F. Malle, M. Scheutz, J. Forlizzi, and J. Voiklis, "Which robot am i thinking about?: The impact of action and appearance on people's evaluations of a moral robot," in *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 2016, pp. 125–132.
- [13] A. Abubshait and E. Wiese, "You look human, but act like a machine: Agent appearance and behavior modulate different aspects of human-robot interaction," *Frontiers in psychology*, vol. 8, 2017.
- [14] M. Salem, G. Lakatos, F. Amirabdollahian, and K. Dautenhahn, "Would you trust a (faulty) robot? effects of error, task type and personality on human-robot cooperation and trust," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015, pp. 141 – 148.
- [15] A. Prakash and W. A. Rogers, "Why some humanoid faces are perceived more positively than others: effects of human-likeness and task," *International journal of social robotics*, vol. 7, no. 2, pp. 309–331, 2015.
- [16] P. J. Hinds, T. L. Roberts, and H. Jones, "Whose job is it anyway? a study of human-robot interaction in a collaborative task," *Human-Computer Interaction*, vol. 19, no. 1, pp. 151–181, 2004.
- [17] Questionnaires, "<https://tinyurl.com/ybw96jgm>."
- [18] Softbank, "<https://www.softbankrobotics.com/emea/fr/robots/pepper>."
- [19] IRobot, "<https://www.irobot.co.uk/>."
- [20] Rethink, "<https://www.rethinkrobotics.com/sawyer/>."