The potential of physical motion cues: Changing people's perception of robots' performance

Pedro Garcia Garcia, Enrico Costanza, Sarvapali D. Ramchurn and Jhim Kiel M. Verame
University of Southampton
Southampton, United Kingdom
{pgg1g14, e.costanza, sdr1, j.verame} @soton.ac.uk

ABSTRACT
Autonomous robotic systems can automatically perform actions on behalf of users in the domestic environment to help people in their daily activities. Such systems aim to reduce users' cognitive and physical workload, and improve well-being. While the benefits of these systems are clear, recent studies suggest that users may misconstrue their performance of tasks. We see an opportunity in designing interaction techniques that improve how users perceive the performance of such systems. We report two lab studies (N=16 each) designed to investigate whether showing physical motion, which is showing the process of a system through movement (that is intrinsic to the system's task), of an autonomous system as it completes its task, affects how users perceive its performance. To ensure our studies are ecologically valid and to motivate participants to provide thoughtful responses we adopted consensus-oriented financial incentives. Our results suggest that physical presence does yield higher performance ratings.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous.

Author Keywords
Visualisation of Automation; Cognitive Psychology; Automated Systems; Robots; User Experience; Perception

INTRODUCTION
There is a growing number of systems able to *automatically* perform actions on behalf of users. Such systems are becoming increasingly widespread in the domestic environment to help people in their daily activities, such as water sprinkling[1] and vacuum cleaning. Moreover, as the domestic environment becomes increasingly instrumented with smart sensors through the Internet of Things (IoT), autonomous systems will be essential to manage the wealth of data such sensors generate, relieving their users of significant cognitive and physical workload involved in performing their daily activities.

While the potential benefits of automatic systems are clear, there are open questions around how users would perceive such systems and their operation [16]. Recently, researchers have investigated the usefulness of existing IoT products, such as the Nest thermostat [25] and vacuum cleaning robots (e.g. iRobot's Roomba) [20, 12]. Results from these studies suggest that because these systems operate autonomously, users do not normally attend to them while they undertake their tasks, so there could be a mismatch between the system performance and users' perception of it. As such, we see an opportunity in designing interaction techniques that may improve how users perceive the performance of automatic systems. In particular, we focus on *notifications*: notification systems are often necessary to alert users when the autonomous operation has been completed (given that users may not attend to it). Is it possible, then, to engineer notifications generated by autonomous system to influence users perception of the system performance?

Against this background, in this paper we report on two lab studies designed to investigate whether showing in person the *physical motion* of an autonomous system as it completes its task can affect how users perceive its performance. Physical motion refers to the robot's movement as it processes its task, hence the motion we refer to is intrinsic to the system's task (this is in contrast to "physical motion" as being independent of task execution). Consensus-oriented financial incentives were used to increase the ecological validity of the studies[4] and motivate our participants to provide thoughtful responses. In particular, the first user study (N=16) focused on comparing two situations: (i) a moving robot in the process of docking and (ii) a static robot that has already completed its task. The aim was to see whether motion can positively change people's perception of the performance of an autonomous robot. Our results demonstrate that this is indeed the case: our participants almost unanimously rated the performance of the moving robot higher compared to a non-moving robot. In the second user study (N=16), we instead focused on investigating whether seeing the motion in person or through a video feed makes a difference. The results suggest it does: physical presence yields higher performance ratings. Indeed, the findings

[1] `http://tinyurl.com/kzk9uuf`
[1] `http://tinyurl.com/legw4zt`

UbiComp '16, September 12 - 16, 2016, Heidelberg, Germany

ACM 978-1-4503-4461-6/16/09…$15.00
DOI: http://dx.doi.org/10.1145/2971648.2971697

of this paper provide implications for how the feedback of autonomous systems can be enhanced to support how people perceive the performance of such systems.

RELATED WORK

Our research aims to evaluate how visual cues can change people's perception about the performance of autonomous robots. As such, we are building upon prior research that has studied transparency for the intelligibility of robotic systems, perception of motion in robots and interactive artefacts, and perception of motion in robots through video and animation. We next elaborate on the literature in these three key areas in turn.

Transparency for the Intelligibility of Robots

Robots are expected to become part of people's everyday life in homes and public areas (e.g. in hotels, trade shows, workplaces, museums) [19, 11]. Therefore, robots should be transparent about their decisions and actions so that people would feel that they understand their behaviour [13].

Kim et al. [13] examined whether different levels of transparency have an effect on people's judgement of blaming an autonomous robot or someone else at the moment the robot presents an unexpected behaviour in a cooperative scenario. In more detail, the robot delivered assemblies of toy pieces that participants place in a tray that the robot had. In particular, a highly transparent robot provided audible feedback about its status. However, they do not focus on how people perceive the performance of the robot with different levels of transparency. Boyce et al. [1] implemented an external interface (screen display) to make the operation of the robot more transparent. Their results showed that increasing transparency can help users understand a robot's environmental conditions and status. In contrast to both of these studies, we do not enhance the existing structure of robots. Instead we utilise their current setup as a way to keep the design of the robots as simple as possible.

Perception of motion in robots and interactive artefacts

Prior studies in HCI, HRI and UbiComp have examined whether people can infer intentionality, emotions or be motivated to interact with robots or artefacts through the visualisation of motion [15, 17, 10, 9, 2, 3]. Instead, in our study, we use the motion of a robot as a visual cue to change how people perceive the robot's performance. Closer to our work, Hoffman et al. [6] conducted a study where an anthropomorphic robot, *Travis*, was used as a speaker dock and music listening companion. Participants observed, listened and evaluated songs played by Travis. For some participants, Travis moved on-beat with the songs played. In contrast, other participants interacted with a moving Travis that was off-beat with the songs. The rest of the participants were introduced to a static Travis. Their results showed that participants rated songs significantly higher when the robot is moving on-beat with the songs than when it is static. Indeed, they pointed out the role of "personal robots as contributors to, and possibly amplifiers of, people's own evaluation of external events" [6].

These findings focus on the evaluation of events that are external to the system e.g. asking people whether they enjoy what they hear, instead of asking them about the quality of the sound produced by the system. Moreover, this work is about entertainment applications, while we look more at mundane or practical applications. In particular, we focus on how people evaluate the performance of such systems. Moreover, they centered their research on anthropomorphic robots. Instead, we are particularly interested with everyday systems (e.g. systems that are used in everyday situations such as cleaning or cooking robots). This is because it may not be practical to modify everyday systems to be anthropomorphic. We intend to focus on maintaining the simplicity of such systems.

Perception of robots motion through video and animation

Previous studies showed how people perceive robots through their physical movement [6]. However, there are other alternatives to interact with robots, such as, videos and animations. Such modalities allow people to visualise robots remotely without having a physical interaction with the system. Takayama et al. [21] examined people's perception of virtual animated robots through a lab study. For the study, the robot covered a variety of activities, such as opening a room, delivering a drink, requesting help from a person to plug into an outlet, and ushering a person into a room. Their results suggest that people are positively influenced by animations showing the outcome of a robot and more specifically that they read robot behaviour with more certainty. However, while the focus of their study is on training we are interested in real time interaction with robots. Additionally while their work is based on virtual animated robots, ours use physical ones. Wainer et al. [24] ran a study with participants that interacted in a collaborative task with an embodiment robot v.s. non-embodiment robot (e.g. simulated and video). In more detail, participants resolve a Towers of Hanoi puzzle following the instructions of the robots. Their results suggest that people perceive an embodiment robot more helpful and enjoyable in comparison with a non-embodiment robot. However, they did not analyse whether people perceive that one type of robot works better than the other, which we present as our key contribution.

MOTIVATION

A large body of work from cognitive psychology investigated how motion and other sensory cues influence our perception of the world. We started from this work to design notification mechanisms that could influence people's perception of autonomous robots.

In psychology "perception" is defined as the process that people follow to identify, interpret, and understand their environment, with the support of sensory (i.e. physical) and cognitive cues (referred to as *high-level of knowledge*) that the nervous system processes [18]. Studies have shown that humans can extract high level information from very basic motion cues [8]. However, in some cases, physical cues are insufficient for the brain to interpret the environment. Hence, the brain uses existing knowledge as a way to make sense of sensory signals (e.g. sight) [5].

Our perception of the world is sometimes influenced by more than one sensory channel. For example, McGurk and MacDonald demonstrated that speech perception is influenced by

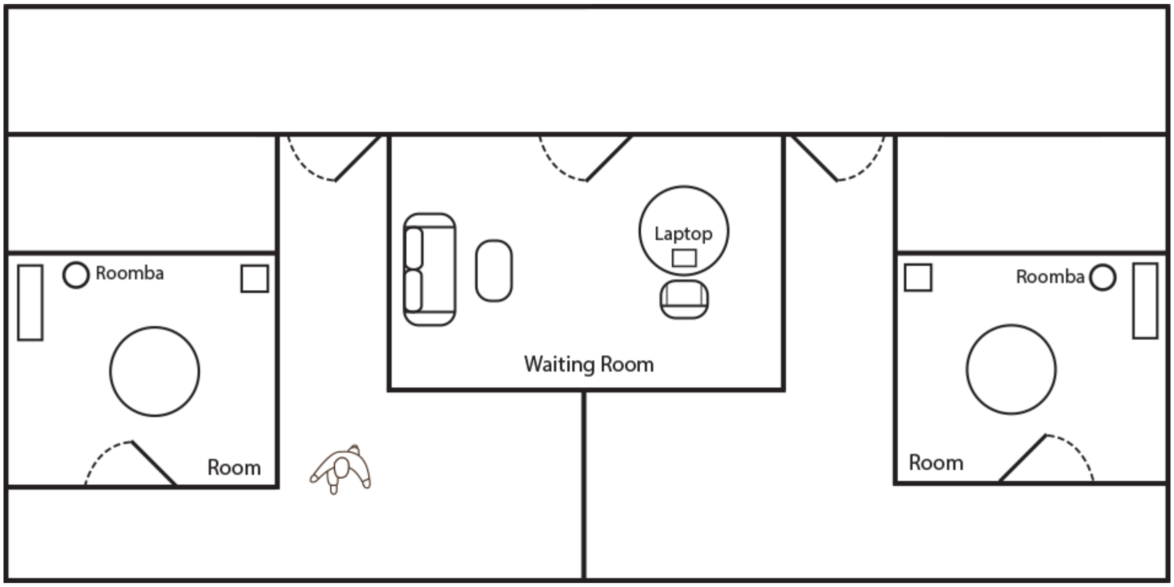

Figure 1. This figure shows a layout of the rooms where we conducted the experiment.

both sound and vision [14]. Vines et al. [23] reported a study where participants rated how much they liked audiovisual clips of clarinet players to investigate how different visual cues affect people's evaluation of the musicians' performance. They found that participants gave a lower score to a clarinettist who did not move compared to a clarinettist with more expressive body motion. This result suggests that an appropriate visual cue can improve people's rating of a non-visual property. Building on such a corpus, we set out to explore whether motion can be leveraged to influence people's perception of autonomous robots.

STUDY DESIGN

A user study was designed and conducted to analyse the effect of motion, as a visual cue, on people's perception. Specifically, the study was designed to test the following hypothesis:

> H1 – The *visualisation of automation*, which is showing system process through motion (that is intrinsic to the system), is an effective visual cue that can positively change people's perception of the system.

Experiment 1

Experiment Design

For this experiment, we selected the Roomba robot because it is an off-the-shelf product designed for domestic everyday use. A within-groups design was used to compare the effect of the visual cue and its absence on the same set of participants, where participants evaluated and compared the performance of two Roomba robots in vacuuming the carpets in two rooms where each was located. Two conditions were defined in our experiment: *no-motion* and *motion*. In the no-motion condition (the control condition), participants saw the robot after it completed its task, having already returned to its charging base. In contrast, in the motion condition (the 'treatment' condition) participants saw the robot moving as it docked in its charging base, having completed its cleaning duties. This movement is the *visual cue* at the centre of our study. In practical terms, the motion condition was implemented through a Wizard of Oz approach whereby an experimenter activated the robot seconds before participants arrived. The study was fully counterbalanced: half of the participants saw first the robot in the motion condition, vice versa for the other half. Moreover, the rooms and the robots were also alternated and fully counterbalanced: half of the participants saw the robot in the motion condition in Room A and the other half saw the robot for the same condition in Room B. The first robot participants saw was referred to as simply 'robot A', and the other one 'robot B', regardless of the condition (so that the naming would not influence the results).

At the beginning of each experiment, participants were told that the task was to compare two different algorithms implemented on each of the two robots. After this introduction, participants were asked to visit two rooms and were given a questionnaire asking them to evaluate on a 5-point likert scale the *cleanliness of the carpet* (from "1 - dirty" to "5 - clean without any chance to improve"). They were then asked to move to a different room to wait until the Roombas finished vacuuming the carpets. Participants were explained that they had to wait in a different room because the algorithms were still work-in-progress so we did not want their judgement to be influenced by their trajectories. Figure 1 shows the layout of the room. While waiting, participants were asked to play a puzzle game [2] on a 13" screen laptop. When the two robots had completed their tasks, participants received a text-based notification shown on the laptop indicating that they could go and evaluate the performance of the Roombas. After receiving the notification, participants visited both rooms one after the other. As described above, in one room they found the robot

[2] http://tinyurl.com/krx3w73

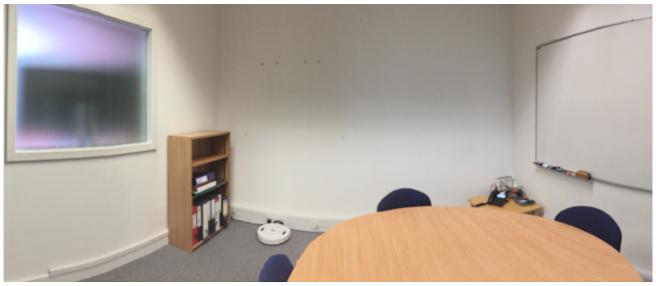

Figure 2. On this figure, we can see one of the rooms where the experiment took place.

already docked, while in the other they saw it docking. After participants had seen the robot docking, we told them that the robot's action of docking was not related to the robot's task of cleaning the carpet. As such, they were allowed to see this part of the robot's process.

After visiting each room, they were asked to continue the questionnaire and evaluate whether there is an *improvement with the cleanliness of the carpet* on a 5-point likert scale (from "stayed the same, did not have an improvement" to "better than before"). The post-task question was phrased differently from the pre-task questions, so the answers cannot be directly compared.

Once they evaluated both rooms, participants were asked to compare the performance of the Roombas. To try and ensure that participants would provide a significant and thoughtful evaluation, we designed a performance-based reward mechanism. Participants were asked which of the Roombas they thought most people would select as the one with the best performance (including the option that both had the same level of performance), and they were told that only if they selected the most popular choice at the end of the study (after we collected data from all participants) they would be rewarded with a £10 voucher (hereafter referred to as *reward-based question*). To check whether participants subjective judgement of the Roombas differed from what they expected the majority of people to choose, after they answered the first question they were presented with a second question, asking them which robot they personally consider to be the one with the best performance, regardless of other people's opinion. This second question (referred later as *non reward-based question*) had no effect on the reward received by the participants.

External validity was a key factor in the design of the experiment to test the effectiveness of visual cues in people's perception when they evaluate the performance of the robots. Therefore, we were particularly careful in keeping a number of variables that could affect participants' perception of the performance of the Roombas constant. These variables were determined through pilot studies:

- Cleanliness of carpets: The Roombas did not actually clean the carpets during the experiment.
- Robot's environment: The rooms used in the experiment were similar to maintain the same conditions (see Figure 2). Moreover, the robots were switched between the two rooms to maintain a fully counterbalanced study design.
- Roombas' task completion time: Both of the Roombas were simulated to vacuum the rooms in 10 minutes and were working simultaneously.
- Robot's model: The two robots used in the experiments were of the model iRobot Roomba 500.
- Evaluation time: Participants were only allowed 15 seconds to evaluate the carpets in each room. This was done to avoid participants spending more time in one room than the other.

Participants

A total of 16 participants (12 female, 4 male) took part in the study and 15 of these were members of the university: PhD and Masters students, none of which had technical background (e.g., not from Computer Science or Engineering). One participant was a homemaker. The ages of these participants ranged from 24 to 53 years old (M = 32.00, SD = 7.46).

Results

Selection of robot with best performance. For the *reward-based* question, 15 out of the 16 participants selected the moving robot (motion condition) as the one with the best performance. The remaining participant selected the robot in the no-motion condition as the best performing one, while nobody indicated that the robots had the same level of performance. For the *non reward-based* question, only one participant expressed a different opinion from that of the previous question, saying that both robots had the same performance. In total, 14 participants considered the moving robot as the better performing robot when answering the non reward-based question. These results are illustrated in Figure 4.

Cleanliness of the carpets. A Mann-Withney test revealed a statistically significant effect ($U = 67.50, p < .05, r = .41$) of the motion on the rating of how clean the rooms were after the operations of the robots. The room in the motion condition was rated on average as cleaner ($mdn = 2.5$) than the room in the no-motion condition ($mdn = 1.5$). Figure 3 shows the means comparison of the two groups. No statistically

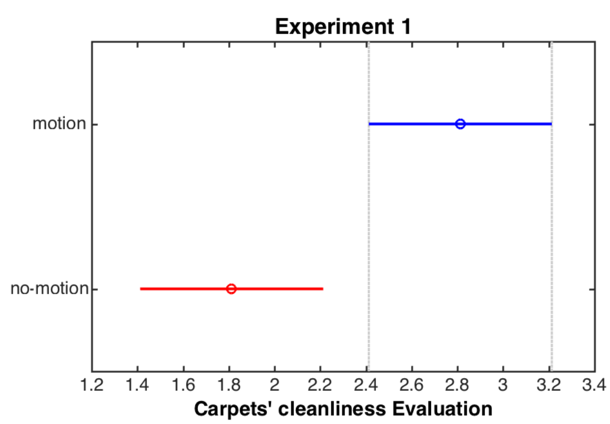

Figure 3. Comparison of evaluation means for rooms' carpet after robots clean, with 95% confidence confidence bars.

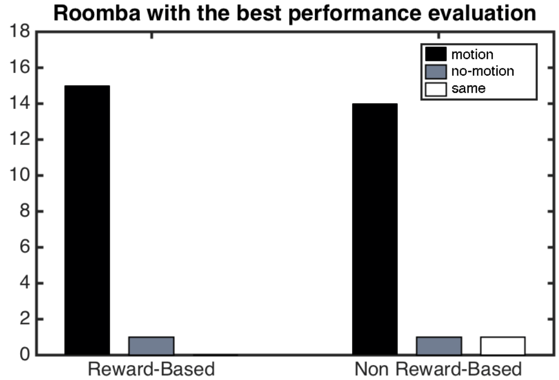

Figure 4. Comparison between participant's evaluation in the reward-based and non reward-based questions

significant differences were found on the ratings of how clean the rooms were before the operations of the robots between the conditions.

Discussion

The results of Experiment 1 confirms H1: motion can be used to change people's perception of how well an automated or automatic system works. The data shows clearly that the change is in the positive direction: all except 2 participants agreed that the robot in the motion condition was the one with the best performance. All except 1 declared that they thought the moving robot would be considered by other participants as the one with the best performance. This finding is further confirmed by the ratings that participants expressed through the likert scales. In the motion condition the room was rated as cleaner than in the no-motion condition, after the robots' operations. As expected, no differences were found on the rating of the rooms before the robots' operations. These results show that the visualisation of automation can be used as a tool to change people's perception of the performance of automated or automatic systems, even in the case of systems that are not anthropomorphic, extending what was previously reported in the literature [6].

Experiment 2

The results from the first experiment clearly show that seeing a robot moving had an effect on our participants' perception of its performance. However, it as far as I know noted that the movement was seen *in person*. Could the same effect be observed if the movement of the robot is experienced through a video feed? Indeed there might be situations in which users are unable to directly see the movement of a robot. To answer this question we designed a second experiment to compare how people perceive the performance of a Roomba when people watch a video of it docking in comparison to watching a Roomba docking in person. As such, we defined a new hypothesis:

> H2 – *Physical visual cues* are more effective than *video-based* cues at positively influencing how people evaluate an autonomous robot that show such cues.

Experiment Design

The design of experiment 2 is the same as experiment 1, except that the *no-motion* condition was replaced by a *video* condition. When participants received the notification that the robot had completed its task in the video condition, they were presented a video showing the Roomba docking. This is similar to the motion condition, but mediated over a video, rather physically seen in the same environment. In this new condition, a video of a Roomba docking was displayed on the laptop computer where the participants played the video game (cfr Experiment 1), and this served as a notification that the Roomba has completed its operation, rather than the text-based notification. For practicality the video was a pre-recorded clip, but it was presented to participants as a live feed from the room (the two rooms had no external windows, making such mockup realistic). Moreover, to avoid details on the video that can change people's perception we used a VGA resolution. To guarantee that participants associate the video-based notification with the correct Roomba the notification was presented to the participants before they visited the room. To accomplish this, the research investigator carried the laptop throughout the duration of the study, including when the rooms are about to be evaluated. Before entering the rooms where the Roombas were, the investigator would show the laptop's screen. For the video condition, this means that they would see the video-based notification right before they enter the corresponding room, therefore guaranteeing that they associate the video with the correct Roomba.

We included some new questions in the final questionnaire. In addition to the *reward-based* and *non-reward-based* questions, participants were also asked why the robot they selected performed better than the other, with a view to understand the motivation behind their choices. Moreover, they were asked whether they would prefer watching a video of the Roomba working or watching the Roomba physically finishing its task and why.

As in experiment 1, we were particularly careful in keeping a number of variables that could affect participants' perception of the system performance constant. These variables were the same as those listed for experiment 1.

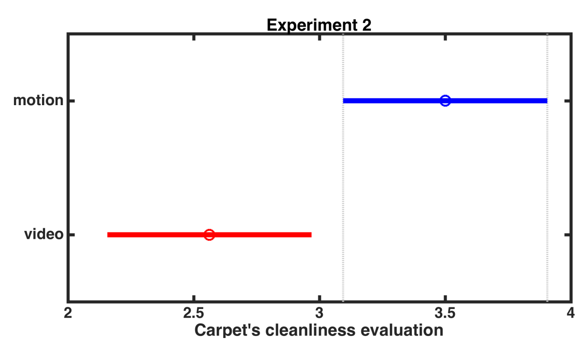

Figure 5. Comparison of evaluation means for rooms' carpet after robots clean, with 95% confidence confidence bars.

Participants

A total of 16 participants (10 male, 6 female) took part in the study and all of them were members of the university: undergraduate and postgraduate students, including a wide range, from Computer Science, English literature, Mechanics, Economics and Psychology students. The ages of the participants ranged from 19 to 37 years old (M = 22.00, SD = 4.39).

Results

Selection of Roomba with the best performance. In total 13 of the 16 participants considered that the Roomba in the motion condition performed better than the Roomba in the video condition. The remaining three participants indicated that the Roomba in the video condition performed the best, while nobody suggested that both robots had the same performance. All participants answered in the same way the *reward-based* and *non reward-based* questions, i.e. they all believed their answer would be the most popular one. These results are illustrated in Figure 7.

Reasons for choosing one Roomba over the other. The responses to the question about why participants selected a particular Roomba as the one performing the best were summarized through open coding. Each response was associated to one or two codes, with five codes used in total: *details, relative, generic, room features,* and *clean already*. Figure 6 illustrates the frequencies of these codes for those who preferred the motion condition and those who preferred the video condition. The code *details* was associated to responses which referred to specific issues in the room, such as "crumbs which lie close to chair legs" and "coffee stains." The code *relative* was used when the responses referred to the comparison of how clean the room was before and after the operation of the robots, such as: "Found the room cleaned by Roomba A much cleaner than it was initially" and "biggest change in cleanness". Comments coded as *generic* included "cleaned the room better" and "The carpet of room B was cleaner than room A". The code *room features* was used when participants referred to the influence of room features on the performance of the robots, such as "less corners for roomba to have difficulty with" and "It seemed to clean tighter spaces better". Finally, one participant stated that the room was clean to start with ("Because the Room B is clean already so it is hard to evaluate the Roomba B performance") so this response was coded as *clean already*.

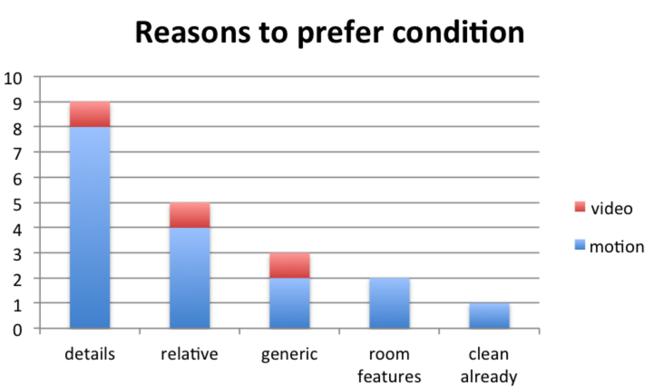

Figure 6. Reasons expressed by participants for preferring one Roomba over the other in Experiment 2.

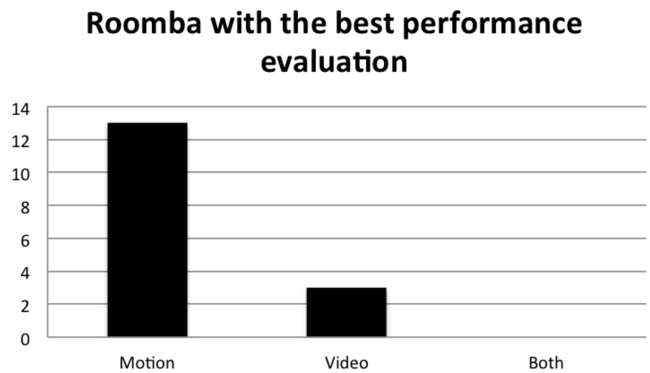

Figure 7. Comparison between participant's evaluation of the two different modalities. Note that all participants responded to the reward and non-reward based questions in the same way.

Cleanliness of the carpets. A Mann-Withney test revealed a statistically significant effect ($U = 70, p < .05, r = .39$) of motion and video on the rating of how clean the rooms were after the robots' operations. The room was rated on average as cleaner in the the motion condition ($mdn = 3$) than in the video condition ($mdn = 2$). Figure 5 shows the means comparison of the two groups.

Modality preference. Ten participants preferred the video over seeing the robot physically move; four participants preferred seeing the robot in person; while the remaining two participants did not have a preference for how they see the robot.

Reason for preferring a modality. The responses to the question about why participants selected one modality over the other were summarized through open coding. Each response was associated to one code, with six codes used in total: *better understanding, convenience, emotional, generic, reliable* and *subjective*. Figure 8 illustrates the frequencies of these codes. An example in the *better understanding* category included "I can understand which part of the room have been cleaned". The *convenience* category included "I do not have to be there till the end", "Can observe the room situation remotely", and "This will save our time while we are doing some other work during the time Roomba was doing its task...". Comments categorised as *emotional* included "fun" and "...physical presence has a more personal effect". Comments in the *generic* category included "You can see the Roomba working physically and the video is helpful" and "Able to see functionality of the roombas". An example comment in category *reliable* category included "On the video you can't see what is happening". Finally, the *subjective* category included "I'm personally a visual person so it illustrates it much better...".

Discussion

The results of Experiment 2 suggest that seeing the robot moving in person positively influences the perception of its performance, compared to seeing it through video. All except

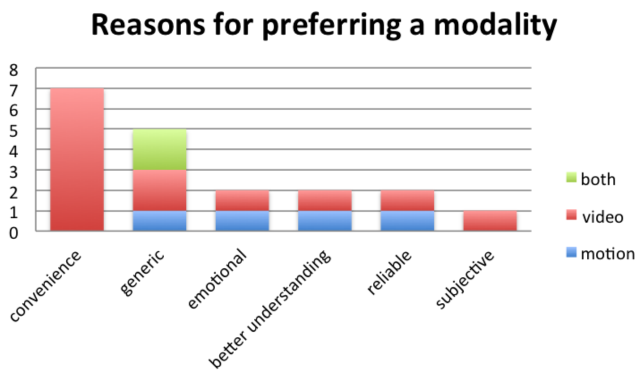

Figure 8. Reasons expressed by participants for preferring one notification modality over the other in Experiment 2.

3 participants agreed that the robot in the motion condition was the one with the best performance. This finding confirms our hypothesis H2. Our participants' ratings of the cleanliness of the rooms further confirm such result: in the motion condition the room was rated as cleaner than in the video condition, after the robots' operations.

The qualitative data about why participants selected one specific robot as the one performing better provide further evidence of the effect of the two different notifications and their potential to influence people's perception of the robots. As illustrated in Figure 6, participants provided generic answers to this question only in three instances. In contrast, in the majority of cases our participants' answers included specific and tangible reasons to support their choice, despite the fact that the Roombas did not actually clean either of the two rooms.

Even though the performance ratings clearly indicate that the Roomba in the motion condition is the most successful one, when participants were asked about their general preference regarding the modality most of them chose the video. The most frequent reason to support this choice was convenience. Such contrast between performance ratings and general preference seems to suggest that participants were not aware of the bias that the motion condition caused on their performance rating. It should also be noted that while the performance rating was related to a financial incentive the question about general preference was not. Therefore it is also possible that participants answered the latter more casually.

GENERAL DISCUSSION AND IMPLICATIONS

The results of both of our experiments indicate that the feedback delivered with notifications can have a considerable influence on people's perception of the performance of autonomous robotic systems. In particular, seeing the robot moving as it finishes its operation in person led our participants to rate its performance higher than not seeing any motion, or seeing the same motion over video.

Compared to prior work [6, 22], the type of motion displayed in our study is very simple, making it very easy and cost effective to take advantage of our findings in existing designs. Indeed, in the case of the Roomba, the motion cue we studied is simply part of the standard operation of the robot. However, additional measures may need to be put in place to drive the user's attention to the motion. For example, presence or location sensing (including e.g. smartphone apps to detect the user's location) may be employed to activate the system when users are physically close to them, or on their trajectory home, leveraging prior work on pattern recognition on GPS traces [7].

These results could potentially apply to a wide range of devices. In the domestic context, smart appliances such as vacuum cleaning robots, washing machines (e.g., seeing the spin cycle confirms the clothes will be clean and dry) and dishwashers (e.g., hearing the dishwasher rinse and shut down confirms all dishes have been cleaned) could be timed according to GPS traces such that when they detect (or predict) that the owners are nearby, they would finish their cycles [7]. A similar approach could also be used for prototyping machines, such as 3D printers, laser cutters and CNC machines.

In addition, the results of our experiments highlight new research opportunities around different ways to present visual cues as new forms of feedback for autonomous robots. While our results, even though on a small sample, show a clear effect, they also open a number of new research questions, for example: Is this effect long lasting? Does it apply to any kind of robots, or even other ubicomp (non-robotic) systems? Is the timing of the cues that are presented important? We believe that the effect we observed may even influence people's inclination to adopt such systems: more research is required in such a direction.

CONCLUSION

In this paper, we have presented two lab experiments, each with 16 participants, designed to investigate whether seeing the motion of autonomous robots in person can positively change people's perception about the performance of the robots. Indeed, our findings suggest that people's perception of the performance of an autonomous robot can be improved for the better through showing them the robot moving, in such a way that they would see it in person. Showing the motion of autonomous systems acts as a visual cue to help people perceive the performance of such systems correctly. In contrast to previous work, our results apply to systems which are not anthropomorphic, hence, the implications can be relevant to a large number of systems. Therefore, we hope that the results presented in this paper will stimulate designers to integrate motion in the feedback of their systems, and researchers to further explore this area.

ACKNOWLEDGMENTS

This work was supported in part by CONACyT, SICYT Morelos, and EPSRC ORCHID and A-IoT, and approved by University of Southampton Ethics Committee (ref: 17155). The data referred to in this paper can be found at http://dx.doi.org/10.5258/SOTON/397976

REFERENCES

1. Michael W. Boyce, Jessie Y.C. Chen, Anthony R. Selkowitz, and Shan G. Lakhmani. 2015. Effects of

Agent Transparency on Operator Trust. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts (HRI'15 Extended Abstracts)*. ACM, New York, NY, USA, 179–180. DOI: http://dx.doi.org/10.1145/2701973.2702059

2. Mason Bretan, Guy Hoffman, and Gil Weinberg. 2015. Emotionally expressive dynamic physical behaviors in robots. *International Journal of Human-Computer Studies* 78 (2015), 1 – 16. DOI: http://dx.doi.org/10.1016/j.ijhcs.2015.01.006
3. Anca D. Dragan, Shira Bauman, Jodi Forlizzi, and Siddhartha S. Srinivasa. 2015. Effects of Robot Motion on Human-Robot Collaboration. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*. ACM, New York, NY, USA, 51–58. DOI: http://dx.doi.org/10.1145/2696454.2696473
4. Doris A Fuchs. 2003. *An institutional basis for environmental stewardship*. Vol. 35. Springer Science & Business Media.
5. V Gregory. 1998. Eye amd Brain, the Psychology of Seeing. (1998).
6. Guy Hoffman and Keinan Vanunu. 2013. Effects of Robotic Companionship on Music Enjoyment and Agent Perception. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction (HRI '13)*. IEEE Press, Piscataway, NJ, USA, 317–324. http://dl.acm.org/citation.cfm?id=2447556.2447674
7. Eric Horvitz and John Krumm. 2012. Some Help on the Way: Opportunistic Routing Under Uncertainty. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 371–380. DOI: http://dx.doi.org/10.1145/2370216.2370273
8. Gunnar Johansson. 1973. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics* 14, 2 (1973), 201–211. DOI: http://dx.doi.org/10.3758/BF03212378
9. Jinyung Jung, Seok-Hyung Bae, and Myung-Suk Kim. 2013a. Three Case Studies of with Moving Products. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 509–518. DOI:http://dx.doi.org/10.1145/2493432.2493442
10. Jinyung Jung, Seok-Hyung Bae, Joon Hyub Lee, and Myung-Suk Kim. 2013b. Make It Move: A Movement Design Method of Simple Standing Products Based on Systematic Mapping of Torso Movements & Product Messages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1279–1288. DOI: http://dx.doi.org/10.1145/2470654.2466168
11. Frédéric Kaplan. 2005. Everyday Robotics: Robots As Everyday Objects. In *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies (sOc-EUSAI '05)*. ACM, New York, NY, USA, 59–64. DOI: http://dx.doi.org/10.1145/1107548.1107570
12. Jun Kato, Daisuke Sakamoto, Takeo Igarashi, and Masataka Goto. 2014. Sharedo: To-do List Interface for Human-agent Task Sharing. In *Proceedings of the Second International Conference on Human-agent Interaction (HAI '14)*. ACM, New York, NY, USA, 345–351. DOI: http://dx.doi.org/10.1145/2658861.2658894
13. Taemie Kim and P. Hinds. 2006. Who Should I Blame? Effects of Autonomy and Transparency on Attributions in Human-Robot Interaction. In *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*. 80–85. DOI: http://dx.doi.org/10.1109/ROMAN.2006.314398
14. Harry McGurk and John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264, 5588 (23 12 1976), 746–748. http://dx.doi.org/10.1038/264746a0
15. Ditte Hvas Mortensen, Sam Hepworth, Kirstine Berg, and Marianne Graves Petersen. 2012. "It's in Love with You": Communicating Status and Preference with Simple Product Movements. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. ACM, New York, NY, USA, 61–70. DOI: http://dx.doi.org/10.1145/2212776.2212784
16. Donald A Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books.
17. Diana Nowacka, Nils Y. Hammerla, Chris Elsden, Thomas Plötz, and David Kirk. 2015. Diri - the Actuated Helium Balloon: A Study of Autonomous Behaviour in Interfaces. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 349–360. DOI:http://dx.doi.org/10.1145/2750858.2805825
18. Daniel L. Schacter, James Eric Eich, and Endel Tulving. 1978. Richard Semon's theory of memory. *Journal of Verbal Learning and Verbal Behavior* 17, 6 (1978), 721 – 743. DOI: http://dx.doi.org/10.1016/S0022-5371(78)90443-7
19. D. Schulz, W. Burgard, D. Fox, S. Thrun, and A.B. Cremers. 2000. Web interfaces for mobile robots in public places. *Robotics Automation Magazine, IEEE* 7, 1 (Mar 2000), 48–56. DOI: http://dx.doi.org/10.1109/100.833575
20. J. Y. Sung, R. E. Grinter, H. I. Christensen, and L. Guo. 2008. Housewives or technophiles?: Understanding domestic robot owners. In *Human-Robot Interaction (HRI), 2008 3rd ACM/IEEE International Conference on*. 129–136.

21. Leila Takayama, Doug Dooley, and Wendy Ju. 2011. Expressing Thought: Improving Robot Readability with Animation Principles. In *Proceedings of the 6th International Conference on Human-robot Interaction (HRI '11)*. ACM, New York, NY, USA, 69–76. DOI: http://dx.doi.org/10.1145/1957656.1957674

22. Jo Vermeulen, Kris Luyten, and Karin Coninx. 2013. Intelligibility Required: How to Make Us Look Smart Again. (2013). DOI:http://dx.doi.org/1942/15998

23. Bradley W. Vines, Carol L. Krumhansl, Marcelo M. Wanderley, and Daniel J. Levitin. 2006. Cross-modal interactions in the perception of musical performance. *Cognition* 101, 1 (2006), 80 – 113. DOI: http://dx.doi.org/10.1016/j.cognition.2005.09.003

24. J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric. 2007. Embodiment and Human-Robot Interaction: A Task-Based Perspective. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*. 872–877. DOI:http://dx.doi.org/10.1109/ROMAN.2007.4415207

25. Rayoung Yang and Mark W. Newman. 2013. Learning from a Learning Thermostat: Lessons for Intelligent Systems for the Home. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 93–102. DOI: http://dx.doi.org/10.1145/2493432.2493489