# A Study on Trust in a Robotic Suitcase

**Conference Paper** · November 2016

DOI: 10.1007/978-3-319-47437-3_18



**6 authors**, including:

Sofia Petisca
ISCTE-Instituto Universitário de Lisboa


Patrícia Alves-Oliveira
ISCTE | INESC-ID | Cornell University


Ana Paiva
Instituto Superior Técnico, Inesc-ID

# A Study on Trust in a Robotic Suitcase

Beatriz Quintino Ferreira[1(✉)], Kelly Karipidou[2,3], Filipe Rosa[3], Sofia Petisca[1], Patrícia Alves-Oliveira[4], and Ana Paiva[1]

[1] INESC-ID & Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
beatriz.quintino@tecnico.ulisboa.pt
[2] KTH Royal Institute of Technology, Stockholm, Sweden
[3] Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
[4] INESC-ID and Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL, Lisbon, Portugal

**Abstract.** This work presents a study on human-robot interaction between a prototype of a robotic suitcase – aBag – and people using it. Importantly, for an autonomous robotic suitcase to be successful as a product, people need to trust it. Therefore, a study was performed, where participants used aBag (remotely operated using the Wizard of Oz technique) for carrying their belongings. Two different conditions were created: (1) aBag follows the participant at a close range; (2) aBag follows the participant on a further distance. We expected that participants would trust more aBag when it was following them at a close range, but interestingly participants seemed to trust more when aBag was further away. Also, regardless of the conditions, the level of trust in aBag was significantly higher after the interaction compared to before, bringing positive results to the development of this kind of robotic apparatus.

**Keywords:** Assistive robotics · Human-robot trust · Proxemics

## 1 Introduction

With the fast pace of technology development, a future with robots existing alongside with humans becomes each day more real. Social robotics and human-robot interaction (HRI) studies become crucial for understanding how this can be achieved and bring positive outcomes. One of the fields where the presence of robots is being studied is in assistive robotics [15], where robots can help people overcome their disabilities or enhance their capabilities. But for this cooperative relationship to occur an important bond needs to exist: trust [11].

With this in mind, we tried to ascertain how much one would trust a robotic platform to carry one's personal belongings. Thus, we present a study of a robotic suitcase prototype (aBag, depicted in Fig. 1). Its purpose is to assist people in carrying their belongings and can therefore be labeled as an assistive robot. Having a robotic suitcase, such as aBag, that can carry your belongings and move itself seems to be convenient for almost everyone, as users can benefit from not having to carry the bag's weight and having both hands free for holding or


A. Agah et al. (Eds.): ICSR 2016, LNAI 9979, pp. 179–189, 2016.
DOI: 10.1007/978-3-319-47437-3_18

doing other things. It is plausible to envisage several scenarios in which such an assistive robot would be extremely helpful, specially for elderly people. Recently, the NUA company announced the development of a robot, the NUA Robotic Luggage [7], designed to perform the same task as aBag. From what is available, it seems that this robotic suitcase is an autonomous version of aBag. Thus, the questions we pose in this work and the study performed become even more relevant and opportune.

**Fig. 1.** aBag: front and side views.

Moreover, some studies [5,14] showed that distance (proxemics) is an important factor in HRI. Thus, when measuring trust towards aBag we defined two different types of distance behaviors that aBag could have towards the user: (1) aBag followed the user closely and (2) aBag moved more freely, keeping a further distance to the user. A study was conducted where participants performed both conditions following a within-subjects design. The research questions that underlie this work are the following: *Will users trust their belongings to a robotic suitcase? And will that trust differ depending on aBag's distance-behavior?* Additionally, our study hypotheses are:

*h1: The perceived human-robot trust will be different before interacting with the robotic suitcase aBag and after interacting with it;*

*h2: The perceived human-robot trust is higher for the condition in which aBag follows the user more closely than for the condition in which aBag moves more freely and further away from the user.*

We expected to confirm both hypotheses, thus finding different results for the level of trust before and after the interaction, in order to show the usefulness of aBag, and at the same time we expected users to trust more when aBag was moving closer to them. In spite of the difficulty of using a small sample of participants some interesting findings emerged, those are further discussed.

## 2 Related Work

In assistive robotics valuable work has been conducted, specifically for rehabilitation, for example mobility assistance for disabled or elderly people (e.g. smart walkers). Nevertheless, when it comes to outsourcing this kind of task to a robot, humans hand over a considerable amount of responsibility. Henceforth, it becomes important to trust robots to perform these tasks. Lee and

Moray define trust as *"the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability"* [6], and it is a multi-dimensional concept, inherently complex to measure. Specifically for HRI, evaluating human-robot trust levels has been a challenging and continuously addressed topic [9,16]. In [4] is presented an attempt to quantify the effects of different factors on perceived human-robot trust, showing that robot characteristics (in particular, its performance) influences greatly the perceived trust. Additionally, [2] studied if humans trust a robot to be their partner in a card game scenario. This study showed that although humans trust a robot to partner with them in the card game, those levels of trust are dependent on previous interactions with the same robot. This is what we expect to inform in our hypothesis *h1*, with users assigning higher levels of trust after interacting with aBag. Moreover, in [2] it was found that because trust is a multifaceted construct, it develops and is expressed differently for humans and robots.

To our knowledge, no other study was performed addressing trust in HRI in an assistive scenario in which the robot carries the belongings of a human.

In such a study, proxemics behavior (the distance the robot would have to the user) becomes a relevant variable to take into account. Different factors have been found to affect HRI proxemics, with studies showing for example that the perceived familiarity with a PR2 robot influences personal space, with people standing closer to the robot when it is perceived as more familiar to them [14]. Another study showed that a more mechanical appearance of the robot seemed to allow people to let the robot come to a closer distance (comparing to a more humanoid one) [13]. When addressing personal space in a social interaction, psychology literature defines four primary zones: intimate, personal, social and public [1]. In a study performed in [5], when people were asked to be followed by a robot the majority of participants preferred to position themselves in a way so that the robot would be in their personal space (approximately 0.5 m to 1.2 m). Following this, we would expect that users would also prefer to have aBag with their belongings closer to them, rather than further away. Therefore, for the condition where aBag followed the user closely, the robot moved in the personal zone (0.5 m-1.25 m), and when aBag moved more freely it situated itself more in the public zone (beyond 3.66 m).

## 3 Methodology

### 3.1 The Robot

aBag is a regular suitcase (of size 49 cm $\times$ 34 cm) attached to the chassi of a remotely controlled car (a *Ninco 1/10 Predator MT-10 2.4G RTR*). All the unwanted parts were removed from both the car (its top case) and the suitcase (its wheels and all the plastic and metal components were detached to make it lighter), and the two were fitted together. To make aBag more believable, a hole was made on the bag's front to expose a smartphone's camera mounted on the inside of aBag, in order to disguise as a sensor that the bag would use to track the user (when questioned by the users the false sensor was explained to afford obstacle avoidance and person-following).

### 3.2 Measures

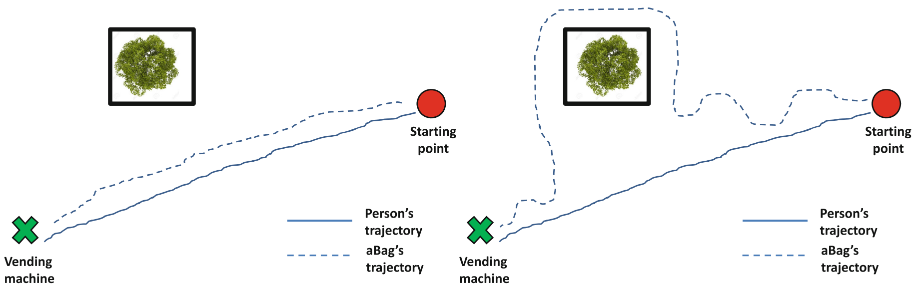
To ascertain how much participants trusted in aBag taking care of their belongings, a trust questionnaire created by K. Schaefer (2013) and specific for HRI was used [11]. This questionnaire was found to be the most appropriate for this study since it is supported by a large body of work and bases the level of perceived trust on the analysis of different dimensions of HRI such as: autonomy, reliability, decision making, failure and communication. These aspects are specially important as the study was performed in an uncontrolled environment. The scale is comprised of 40 items, measuring trust on a 0–100 percentage scale (answers are given in intervals of 10 %) providing a percentage trust score with the sum of each item score. The questionnaire was used before and after the interaction for each of the conditions. Examples of questions from the trust questionnaire are: *"What percentage of the time does this robot follow directions?"* and *"What percentage of the time is this robot reliable?"*.

All interactions with aBag were video recorded and the number of times each participant gazed back at aBag were coded from the videos. Two different coders coded half of the videos that were randomly chosen and a good level of agreement with the Spearman's coefficient $r_s = 0.968$ was obtained, which seemed reasonable to justify only one coder coding the remaining videos. In total, 1.5 h of recorded material was coded.

### 3.3 Pilot Study

Before performing the main study with the final test-users, a pilot study was conducted in order to test the planned experimental procedure. Three volunteers, all male aged 25, 26 and 25 years, performed the experiment at the same place where the main study was planned to take place. Three different conditions for aBag's movement behavior were tested. In the first condition aBag was kept at a small distance from the user, following him/her closely. The second condition was identical to the first but the distance kept between aBag and the user was considerably larger. For the third and final condition, a more "exploring" behavior was applied, in which aBag performed several movements to explore the area freely, as an attempt to make it appear more autonomous. Before and after experimenting with each condition the participants were asked to fill in the trust questionnaire [11]. Out of the three piloted conditions, two were defined (depicted in Fig. 2) for the final study based on the direct feedback from the three pilot study participants and on their answers to the questionnaire:

*Condition 1. aBag follows the user, taking the same trajectory at a small distance behind* - this was the first condition and was chosen because there were indications in the results and observations of the pilot study that this was the behavior the pilot users preferred aBag to have.




(a) Illustration of condition 1. (b) Illustration of condition 2.

**Fig. 2.** A schematic image displaying the difference between the two conditions used for aBag in the main study.

*Condition 2. aBag follows the user more freely, keeping a further distance to the user compared to condition 1, simulating a more autonomous behavior* - this condition became a fusion of the second and the third one, reducing the "exploring" behavior, because of indications that too much exploration was unnatural for aBag since it was aimed to follow the user.

### 3.4 Main Study

**Participants.** A total of 18 university students (11 male and 7 female, aged between 17 and 47 years old, $M = 24$; $SD = 7.0$) performed the experiment. 11 of the participants (8 male, 3 female, $M = 26$; $SD = 8.8$) performed both conditions, so each participant either did first condition 1 and after condition 2, or vice-versa. To counterbalance learning or adaptation effects, the initial condition was randomly chosen for each participant. The remaining 7 participants (3 male, 4 female, $M = 22$; $SD = 3.1$) were only able to perform one of the conditions so they were excluded from the main analysis (results from these participants were only used to investigate the level of trust in aBag regardless of the conditions). All participants signed a consent form in order to be part of the study and authorizing for the sessions to be recorded. The study was a within-subjects design and lasted approximately 30 min. The between-subjects design (for the participants allocated only in one condition) lasted approximately 20 min. At the end of the experiment participants received a thank you gift, which was incorporated in the task when interacting with the robot, see *Experimental Procedure* section.

The study took place in the entrance hall of a Portuguese University, using the Wizard of Oz technique. The wizard was placed on the second floor having a clear sight of the entrance floor without being seen, and kept consistent behaviors for the two conditions using waypoints defined in the environment. The material used was the robotic suitcase (aBag) and a video camera for recording the interaction.

**Experimental Procedure**

(i) The participant was led into a room, where he/she was asked to fill in: a consent form approving the whole experimental procedure and the pre trust questionnaire about how much he/she would trust in aBag (having a picture of the robot to look at).

(ii) Then, the participant was led out to the entrance hall and introduced to aBag. In order to simulate aBag being autonomous, a smartphone was given to the participant and he/she was told that the person wearing that phone would be the person that aBag would follow. However, the phone was not connected in any way to aBag, instead aBag was controlled by a researcher that acted as a wizard.

(iii) After this, the participant was told that the purpose of the study was to simulate a real scenario where a person uses aBag as his/her own robotic suitcase while performing a task. In order to simulate this the participant was asked to put something of value for him/her (wallet, mobile phone) inside aBag, so that aBag could carry it. Then he/she was asked to go to a vending machine, approximately 50 m away, while aBag would follow (see Fig. 3) and buy something he/she likes for one euro (this money was given from us and served as a thank you gift). After buying something, the participant was asked to return back to where he/she started.

(iv) When the participant came back he/she was asked to fill in the post trust questionnaire.

(v) Then, the participant was brought again to perform the same task (step (iii)) with aBag, but now on a different condition.

(vi) At the end of the second interaction, the participant was asked to fill in the post trust questionnaire again and was thanked for his/her collaboration.

The fact that aBag was acting in different ways for the two interactions was not explicitly explained to the participants. The 7 participants that could not perform the whole experiment stopped the procedure at step (iv).[1]

**Fig. 3.** Participant using aBag as his own robotic suitcase.

[1] A video presenting the experiment using the recordings from the participants' interactions with aBag is available at https://www.youtube.com/watch?v=M4mw5WX-AS8\&feature=youtube.

## 4 Results

### Results from the Trust Questionnaires

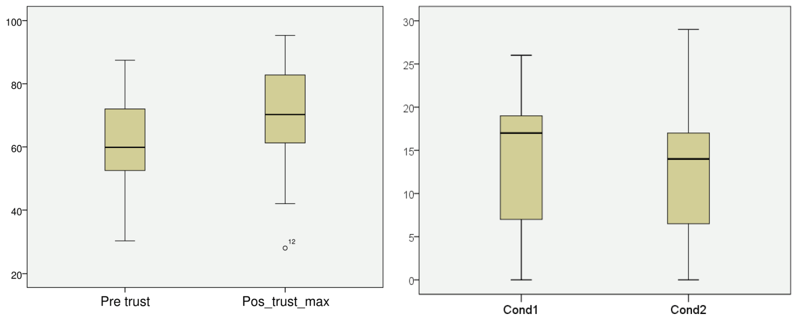
To test hypothesis *h1* "The perceived human-robot trust will be different before interacting with the robotic suitcase aBag and after interacting with it", the data from all the 18 participants was used. For this purpose, the maximum of the post questionnaire from participants that did both conditions were used together with the post values from the participants that performed the experiment only once, to form one score of post interaction trust for each participant. So, for each participant two trust values were gathered, the pre and post trust score, regardless of the condition(s) (see Fig. 4a). When analysing the distribution with a Shapiro-Wilk we found a non normal distribution ($\rho < 0.05$) for the trust scores. Therefore, a nonparametric test for repeated measures was applied to ascertain if there were differences in the trust scores. The Related-Samples Wilcoxon Signed Rank Test confirmed *h1* with $Z = -2.004$, $p = 0.045$. From this result it is possible to conclude that the trust was significantly different before and after interacting with aBag (the median of the pre trust score was 59.88 while for the post trust was 70.25). Hence, the level of human-robot perceived trust is shown to increase after interacting with aBag compared to the participants' measured trust before meeting the robot.

In order to test hypothesis *h2* "The perceived human-robot trust is higher for the condition in which aBag follows the user more closely than for the condition in which aBag moves more freely and further away" a one-way analysis of variance (ANOVA) test was applied using the data from the 11 participants that performed the experiment following a within-subject design. In this case three groups are compared: the questionnaire answers before any interaction, the answers after interacting with aBag when it behaved according to condition 1, and the questionnaire answers after interacting with aBag in condition 2. The results from this test were not significant, which is probably due to the small number of participants. Thus, no significant differences seem to exist in trust between the two conditions. However, there seems to be indications that users trust more in aBag when it behaves according to condition 2 (aBag moving more freely and keeping a further distance to the user), as trust is generally higher after condition 2 (mean and standard deviation for condition 2 are $M = 69.7$, $SD = 17.5$, and for condition 1 $M = 67.1$, $SD = 17.1$). Moreover, the significance values are always lower when comparing the pre and post trust after condition 2 than when comparing the answers for pre and post trust after condition 1.

### Results from the recordings of the interactions with aBag

The video recordings of the experiment were analysed by counting the number of times the 11 participants who performed both conditions looked back at aBag during each one of their interactions. This number might be an indication of how much the participants trusted aBag, assuming that when a person is doubting aBag he/she looks back more often than when he/she trusts it.

We found a non normal distribution ($\rho < 0.05$) for the number of times participants looked at aBag. Hence, a non-parametric Related-Samples Wilcoxon Signed Rank Test was applied to test this data. The results show that there are no statistically significant differences, with $Z = -0.06$, $p = 0.95$, so we retain the null hypothesis. Nonetheless, the data distribution suggests that people tend to look back at aBag slightly more often during condition 1 when aBag is following them closer (the median is higher for condition 1 than for condition 2) (see Fig. 4b). Moreover, the participants who look more often at aBag seem to have a more coherent behavior, as the number of times they look back is more concentrated than the number of times of people who look less. From this observation, it seems that there is a certain number of times/frequency of looking back at the robot that satisfies the users while interacting with aBag.




(a) (b)

**Fig. 4.** Boxplots of (a) the pre and post trust scores, regardless of interaction condition; (b) number of times participants look back, for the two conditions.

When further exploring this quantitative data from the videos, a considerable difference is noted between the number of times male and female participants looked back while performing condition 2. By performing a nonparametric test again, this time the Independent-Samples Mann-Whitney U test, a significant difference in the number of times male and female participants look at aBag, while interacting with it in condition 2, is verified. Yet, this result uses a very small dataset, so it is not possible to make any claim from it. Nevertheless, the data seems to suggest that female participants tend to look more at aBag, probably because women are more used to carry a purse in their daily routines, and consequently look more after it, than men (who usually carry their personal belongings on their pockets). This result may point towards the very interesting aspect of differences in proxemic preferences according to gender [12].

## 5 Discussion

As an attempt to answer the first research question "Do users trust their belongings to aBag?" statistical analysis of the answers to the trust questionnaires confirms hypothesis *h1* as it shows that the perceived human-robot trust is significantly different before and after interacting with aBag. Specifically, the perceived trust was higher after having interacted with aBag than before meeting the robot. This is an interesting result showing that after experimenting with aBag capabilities, participants felt aBag more trustworthy than before interacting. One can argue that a real implementation of a solution to the person-following problem as in [3,10] rather than a Wizard of Oz technique, could have been used to perform this study. Nonetheless, we opted for the latter approach due to the high complexity and uncontrollability of the chosen realistic scenario (with large groups of people passing by), where an automatic algorithm could be dangerous. On the other hand, the novelty effect could also have had a part in this since robotic suitcases are not common yet and people may tend to underestimate their capabilities. Also, the way participants were introduced to aBag before interacting with it might have affected their feedback, as they were only shown a photo of aBag - which conveys minimal knowledge - when filling in the pre questionnaire. However, depite these factors, this presents an interesting result.

Regarding the second question, "Will the users' trust differ depending on aBag's behavior?", no significant statistical results were obtained. Yet, results suggest that trust is higher for condition 2 (the further away and more freely moving condition). Although this indication is opposite to what is formulated in hypothesis *h2*, it is in line with some previous work on proxemics in which users have shown to prefer social robots that keep a higher distance from them and not invading their intimate space [8]. A possible explanation for this could be that when the hypothesis was formulated aBag was thought of as a more machine-like robot and hence shorter distance was expected to be preferred. But, possibly aBag was perceived as more autonomous and "lifelike", having more similarities to the human user rather than an object and thus conveying more "trust" for a greater distance from the user.

The results obtained from the video analysis with the amount of times participants looked back at aBag were not statistically significant. However, there were indications that participants looked less at aBag during condition 2. Therefore, if, in fact the number of times people look at the robot can be considered as a proxy to evaluate trust (less number of looks equivalent to a higher trust), these indications are in agreement with the suggestion from the questionnaires answers that trust tends to be higher for condition 2.

## 6 Conclusions and Future Work

In conclusion, we found that the perceived human-robot trust is significantly higher after the participants have interacted with aBag than before any interaction, independently of study condition, which shows the capabilities of aBag as a robotic suitcase. The two different behaviors (conditions) that were created in order to look for differences in trust in aBag did not present any significant difference, giving only vague indications that participants seemed to trust aBag more when it moved more freely and on a further distance than when it followed them on a closer range.

In the future, it would be interesting to perform a similar study with a larger sample and using target population (e.g. elderly people, wheelchair users). Furthermore, it would be of interest to change the environment where the study is performed, considering creating a more constrained environment where actors may play up the same scenario for each participant without their knowledge, instead of having a dynamical environment as the one in our study.

Finally, it would be interesting to investigate the possible influence in the perceived trust in aBag from more different variables such as age, previous knowledge, experience with robots, and personality factors.

**Acknowledgements.** This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013. P. Alves-Oliveira acknowledges a FCT grant ref. SFRH/BD/110223/2015. The authors are solely responsible for the content of this publication. It does not represent the opinion of the EC, and the EC is not responsible for any use that might be made of data appearing therein.

## References

1. Argyle, M.: Bodily Communication, 2nd edn. Routledge (1975)
2. Correia, F., Alves-Oliveira, P., Maia, N., Ribeiro, T., Petisca, S., Melo, F.S., Paiva, A.: Just follow the suit! trust in human-robot interactions during card game playing. In: RO-MAN (2016, in Press)
3. Gockley, R., Forlizzi, J., Simmons, R.: Natural person-following behavior for social robots. In: HRI (2007)
4. Hancock, P., et al.: A meta-analysis of factors affecting trust in human-robot interaction. Hum. Factors: J. Hum. Factors Ergon. Soc. **53**(5), 517–527 (2011)
5. Huttenrauch, H., et al.: Investigating spatial relationships in human-robot interaction. In: IROS (2006)
6. Lee, J., Moray, N.: Trust, control strategies and allocation of function in human-machine systems. Ergon. **35**(10), 1243–1270 (1992)
7. Libman, A., et al.: NUA – The carry-on that follows you wherever you go. In: NUA (2015). http://unbouncepages.com/nuarobotics/.
8. Pacchierotti, E., et al.: Evaluation of passing distance for social robots. In: IEEE International Symposium on Robot and Human Interactive Communication (2006)
9. Salem, M., et al.: Evaluating trust and safety in HRI: practical issues and ethical challenges. In: Emerging Policy and Ethics of Human-Robot Interaction (2015)
10. Satake, J., Miura, J.: Robust stereo-based person detection and tracking for a person following robot. In: ICRA – Workshop on People Detection and Tracking (2009)
11. Schaefer, K.: The perception and measurement of human-robot trust. Ph.D. Dissertation, University of Central Florida, USA (2013)
12. Syrdal, D., et al.: A personalized robot companion? – the role of individual differences on spatial preferences in HRI scenarios. In: RO-MAN (2007)
13. Syrdal, D., et al.: Sharing spaces with robots in a home scenario-anthropomorphic attributions and their effect on proxemic expectations and evaluations in a live HRI trial. In: AAAI Fall Symposium: AI in Eldercare: New Solutions to Old Problems (2008)

14. Takayama, L., Pantofaru, C.: Influences on proxemic behaviors in human-robot interaction. In: IROS (2009)
15. Tapus, A., et al.: Socially assistive robotics [grand challenges of robotics]. IEEE Robot. Autom. Mag. **14**(1), 35–42 (2007)
16. Can we trust robots? [Special report]. IEEE Spectrum **53**(6), 26–27 (2016). IEEE