CrossMark

The Development of a Scale to Evaluate Trust in Industrial Human-robot Collaboration

George Charalambous[1] · Sarah Fletcher[1] · Philip Webb[2]

Abstract   Trust has been identified as a key element for the successful cooperation between humans and robots. However, little research has been directed at understanding trust development in industrial human-robot collaboration (HRC). With industrial robots becoming increasingly integrated into production lines as a means for enhancing productivity and quality, it will not be long before close proximity industrial HRC becomes a viable concept. Since trust is a multidimensional construct and heavily dependent on the context, it is vital to understand how trust develops when shop floor workers interact with industrial robots. To this end, in this study a trust measurement scale suitable for industrial HRC was developed in two phases. In phase one, an exploratory study was conducted to collect participants' opinions qualitatively. This led to the identification of trust related themes relevant to the industrial context and a related pool of questionnaire items was generated. In the second phase, three human-robot trials were carried out in which the questionnaire items were applied to participants using three different types of industrial robots. The results were statistically analysed to identify the key factors impacting trust and from these generate a trust measurement scale for industrial HRC.

Keywords   Human-robot collaboration · Trust scale · Industrial robot

✉ George Charalambous
george.ch@hotmail.co.uk

[1]   Industrial Ergonomics and Human Factors Group, Centre for Advanced Systems, School of Engineering, Cranfield University, Bedford, UK

[2]   Aerostructure Assembly and Systems Installations Group, Centre for Advanced Systems, School of Engineering, Cranfield University, Bedford, UK

1 Introduction

In this section we present a review of literature to introduce the concept of industrial HRC, the importance of trust in human-robot teams and existing measures of trust to set out the research problem.

1.1 Industrial Human-robot Collaboration

A significant amount of assembly tasks in various manufacturing processes still require the flexibility and adaptability of the human operator [1]. In such processes, it is neither feasible nor cost-effective to introduce full automation. The manufacturing industry has shown growing interest in the concept of industrial robots working as teammates alongside human operators [2–5]. In light of recent technological developments, health and safety regulations have been updated to reflect that in some circumstances it is safe and viable for humans to work more closely with industrial robots [6]. Industrial HRC can enhance manufacturing efficiency and productivity since the weakness of one partner can be complemented by the strengths of the other [7]. However, the integration of humans and robots within the same workspace can be a challenge for the human factors community. For example, the installation of large assemblies requires operators to cooperate with large and high payload robots under minimised physical safeguarding [8]. One key aspect that can determine the success of a HRC system is the degree of trust of the human operator in the robotic teammate [9–11]. With the concept of industrial HRC being embraced further, trust needs to be explored in depth in order to achieve successful acceptance and use of industrial robotic teammates.

1.2 Trust in Automation

The development of trust is essential for the successful operation of any team [12]. In the context of human-automation teaming, trust can influence the willingness of humans to rely on the information obtained by an automated system, particularly in risky and uncertain environments [10,13]. Lack of trust will eventually lead the operator to intervene and take control [14]. Lee and See [15] defined trust as "*the attitude that an agent will help achieve an individual's goals in a situation characterised by uncertainty and vulnerability*" (p. 54). Therefore, the outcomes of the automation will lead to adjustments in the degree of trust in the automated system. Also, the authors identified trust antecedents based on three factors, namely purpose, process and performance. The purpose factor is related to the level of automation used, the process factor relates to whether the automated system employed is suitable for the specific task while the performance factor relates to the system's reliability, predictability, and capability. In addition, the degree of system's transparency and observability available to the human partner has been found important for the development of trust in human-automation interaction [16]. Furthermore, task complexity has been suggested to have an impact on the level to which the human operator relies on the automated system [17,18]. Research has also been directed to investigate people's perceived reliability of automated assistance versus human assistance [19] and machine-like agents versus human-like agents [20]. Dzindolet and colleagues [19] found that humans tend to see the automation as being more reliable compared to a human aid, although the same information was provided both by the automation and the human aid. With increasing risk levels, human reliance on automation support increased when compared to human support. Potentially this can lead to automation misuse or overtrust, which can be detrimental [9,21]. Therefore, calibrating appropriate levels of trust is vital for the success of the interaction.

1.3 Trust in Robots

Although robots encompass a degree of automation, they also possess different attributes not possessed by general automated systems. For instance, robots can be mobile, have different degrees of anthropomorphism and tend to be purpose-built. These attributes introduce a degree of uncertainty not found in general automated systems and for this reason robots need to be studied independetly [22]. Subsequently, trust development in human-robot teams may be different to when humans interact with automated systems. Previous literature has suggested that little research was directed in addressing trust in human-robot interactions [13] while other researchers supported that trust has been assessed in terms of automation and then applied in the domain of human-robot teaming without considering the different attributes related to robots [23]. Various factors have been suggested to influence trust development in human-robot interactions. Hancock et al. [24] carried out a meta-analytic review of 29 empirical studies aiming to quantify the effects of various factors influencing human-robot trust. Their findings highlighted the significance of robot-related factors. Robot related performance-based factors (e.g. reliability, predictability, behaviour) and attribute-based factors (e.g. size, appearance, movement) were found to be of primary importance for the development of trust. Environmental factors (e.g. performance factor, task complexity), were identified to have a moderate influence on trust, while little effect was found from human-related factors. Thus, different robots attributes should be considered when assessing trust. However, industrial robots can be different than social, healthcare and military robots and very little research has been directed towards understanding the development of trust in industrial HRC.

1.4 Measuring Trust

Existing measures of trust have been heavily focussed on automation, such as automated teller machines [25] and automated process control systems [26–28]. However, as discussed earlier, the development of trust in human-robot teams can be different to human-automation interactions [13,23]. A trust measure for human interactions with military robotic systems has been developed by Yagoda and Gillan [23] while, more recently, Schaefer [29] developed a trust scale to evaluate changes in trust between an individual and a robot. Although the aforementioned studies enhance our understanding of trust development in HRI, the context is different. In a military human-robot teaming, the functions of both agents are very different from an industrial scenario. Also, industrial robots come in various shapes, sizes, end-effectors and degrees of anthropomorphism according to the operation being utilised for. Thus a generic trust scale might not be suitable for a purpose-built robot such as the ones used in the industrial environment. Trust development in an industrial robot can potentially be influenced by other context-related factors. To our knowledge, no measure exists which specifically evaluates trust in industrial HRC.

1.5 Research Problem, Aim and Objectives

Although trust has received extensive attention, little research has focused on understanding trust development in industrial HRC. To appropriately understand the development of trust between human workers and industrial robots, it is vital to effectively quantify trust. Such a measurement tool would offer the opportunity to system designers to identify the key system aspects that can be manipulated to optimise trust

in industrial HRC. The aim of this study is to develop an empirically determined psychometric scale to measure trust in industrial HRC. Principal objectives were: (i) Exploratory study: Identify the dimensions of trust relevant to industrial HRC and (ii) Trust scale development: Develop a reliable psychometric scale to measure trust in industrial HRC.

2 Exploratory Study

Due to little understanding regarding the influence of trust in an industrial context an exploratory study was carried out to collect participants' opinions qualitatively. This approach led to the development of trust related themes relevant to the industrial context. Following this, a pool of items was developed describing the identified trust-related themes.

2.1 Method

2.1.1 Participants

21 participants (seven females, 14 males; M = 26.6, SD = 4) were recruited from Cranfield University. 20 participants reported having no prior experience interacting with robots or other form of automation, while one participant reported having used a computer numerically controlled machine.

2.1.2 Design

An exploratory study was performed in laboratory conditions. Participants interacted with two industrial robots (low payload and a medium payload), one at a time, to complete a pick and place task. Because this was an exploratory study it was chosen to give participants the experience of interacting with a smaller robot leading to a bigger one.

2.1.3 Materials

Two types of industrial robots were used as shown in Fig. 1 a small scale robot (payload of 5 kg) and a medium scale robot (payload of 45 kg). The small scale robot has built-in safety. In each condition, the robot picked up and handed to participants two flexible stainless steel industrial pipes approximately 60 cm long. For the interaction with the medium scale robot a laser scanner was used to ensure safe separation between the robot and the participant [6].

2.1.4 Task

The task was identical for both robots. The robot picked up the two industrial pipes, one at a time and brought them to the participant at their standing location. When the robot stopped, participants took hold of the pipe. Then the robot gripper released the pipe. Participants positioned the pipe on a table next to them. Then the robot picked up the second pipe and executes the same task.

2.1.5 Data Collection

Previous research identified that a non-industrial robot's performance-related and attribute-based factors had the highest influence on trust, while environmental related factors had moderate effect [24]. Therefore, a semi-structured interview was chosen. The interview guide is provided in the Appendix 1.

2.1.6 Procedure

Participants were informed regarding their right to withdraw and anonymity and gave their written consent. Participants were initially taken to the small scale robot. The researcher instructed participants regarding the interaction task. Participants observed a short robot demonstration to familiarise themselves with the robot. Then the task was executed. Upon completion of the task, data were collected via a one-on-one interview. Following this, participants were taken to the medium scale robot. An identical procedure as before was followed. Interviews were audio recorded with the partici-

Fig. 1 Materials used for the exploratory study

pant's consent. Interviews took place in the robot cell. No other work was undertaken during the interview to minimise participant disruption. Average interview time was four minutes.

2.1.7 Data Analysis

Interviews were fully transcribed and analysed using the Template Analysis' in accordance with guidelines provided by King [30]. The process involves the development of a coding template representing the major themes identified in a hierarchical form so that top level codes represent broad themes while lower level codes represent sub-themes. Care was given to code themes identified in a small minority of transcripts. The template structure was revised iteratively to ensure it reflected the data in the most suitable manner. Interviews were read thoroughly and phrases were classified into three elements: (i) robot (ii) human and (iii) external. Each of these elements was assigned a letter to assist with the coding procedure (e.g. 'R' for robot element, 'H' for human element). Then, emerging trust-related themes were identified and assigned a unique code number. For example, for the robot element two major themes were identified: (i) robot's performance (R1) and (ii) robot's physical attributes (R2). Following this, each theme was analysed further into lower level themes and a unique letter code was attached. For instance, robot's performance included two lower level themes: (i) robot's motion (R1m) and (ii) robot and gripper reliability (R1r). The derived coding template is shown in appendix 2. An inter-rater reliability was carried out to confirm the level of consensus between raters and, therefore, the suitability of the developed template. Two independent raters were approached to assist with the triangulation process. The coding template was used by both raters individually to code the interview transcripts. Results were tabulated for calculation of the Cohen's kappa statistic. The Cohen's kappa statistic was chosen because it corrects for the probability of agreement by chance thus giving a more conservative result when compared to simple agreement percentage. The Cohen's kappa statistic among the experimenter and the raters were: Experimenter-Rater 1: 0.73; Experimenter-Rater 2: 0.66; Rater 1–Rater 2: 0.68. The average agreement was 0.69.

2.2 Exploratory Study Results

Data analysis revealed that lower-level themes could be grouped in three major elements: robot, human and external. Each of these elements consisted of a number of trust-themes which were then decomposed into lower-level themes. Low-level themes were prioritised on the basis of frequency with which they appeared in the data analysis. This is shown in Table 1.

Table 1 Frequency of trust-related themes

Element	Trust-related themes	Lower level theme	Frequency
Robot	Performance	Motion of the robot	21
		Robot and gripping mechanism reliability	20
	Physical attributes	Robot's size	18
		Robot's appearance	15
Human	Safety	Personal safety	17
		Robot's safety features	11
		Safe programming of the robot	6
	Experience	Priorexperiences with robots	14
		Mental models of robots	9
External	Task	Task complexity	15

2.3 Exploratory Study Discussion

2.3.1 Robot Element

The robots' performance was one the most highly discussed themes among participants. Specifically, the motion of the robots was found to be a key trust-relate topic. All of the participants discussed that their trust was influenced by the way the robots moved and the speed at which they grasped the components. Participants described that the robots employed a smooth and fluid motion when moving which was not found to be uncomfortable. Also, participants highlighted that the speed at which it grasped the components allowed sufficient time to react. Another prevailing attitude among participants was the perceived reliability of the robots and the gripping mechanisms. Participants discussed that they could trust the robots because they completed the respective tasks accurately. Also, participants paid attention to the gripping mechanism of each of the robots. Participants felt they could trust the robot because the gripping mechanism did not drop the components during the collaboration.

Physical attributes received attention by the participants. The majority of participants elaborated that the robots' size had an influence on their perceived trust in the robot upon first encounter. Most of the participants felt intimidated by the size of the medium-scale robot prior to the interaction. The dominant view is that upon encountering the robot, participants felt intimidated by its size and appeared to be worried about interacting with it. Some participants discussed that the robots' general appearance can influence their trust. Participants found that a robot with a simple design is preferred

to interact with as it is perceived less like a robotic machine which increases trust.

2.3.2 Human Element

Safety was among the most frequently discussed themes. 17 participants mentioned that trust in the robots was influenced by their feeling of personal safety during interaction. Participants' comments suggested the main safety concern was to avoid being hit by the robot. Furthermore, 11 individuals discussed that being aware of the robot's safety features (e.g. laser scanner) enhanced their perception of safety, made them more comfortable interacting and increased their trust. Also, six participants elaborated they had faith that the robot had been programmed correctly by its operator.

In addition, prior experiences with robots received attention by participants. 14 participants suggested that any prior experiences interacting with industrial robots would have influenced their trust. Specifically, participants elaborated that having prior exposure to similar robots would have reduced their initial anxiety. In addition, nine participants commented their trust in the robot was affected by their mental models. It appeared that participants had pre-conceived notions of robots, mainly through mainstream movies, and these had an initial influence in their level of trust. Participants mainly discussed how surprised they were with the smooth motion of the robot. Some participants held the belief that industrial robots are monstrous, fast and jerky.

2.3.3 External Element

The complexity of the interactive task was the only external-related trust theme emerging through the interviews. 15 participants discussed that the complexity of the task had an influence on their trust towards the robot. Participants commented that the interactive task was not significantly challenging and this aided them to have greater trust in the robot.

2.3.4 Item Generation

A number of trust-related themes relevant to trust in industrial HRI were identified. A questionnaire was developed with twenty-four items relevant to each low-level theme. The items were developed with the assistance of two members from the department of Human Factors at Cranfield University who are knowledgeable on industrial robots. All items developed and their scoring directions are shown in Appendix 3. Reverse-phrased items are necessary in order to reduce participant response bias. The items were randomly placed in the survey.

3 Trust Scale Development

Three human-robot trials in laboratory conditions were carried out using three different types of robots. Tasks represented potential industrial scenarios where humans and robots would collaborate. Three independent groups of participants were recruited. Upon completing the task, participants completed the survey developed from the exploratory study.

3.1 Method

3.1.1 Participants

In study 160 participants took part (15 female, 45 male; $M = 30.6$, $SD = 9$) from Cranfield University. 19 participants reported having some experience with robots and automation while 41 reported having no prior experience with robots and automation. In study 250 participants took part (13 female, 37 male; $M = 30.9$, $SD = 9.6$) from Loughborough University. 20 participants reported having some experience with robots and automation while 30 reported having no prior experience with robots and automation. In study 345 participants (19 female, 26 male; $M = 30.7$, $SD = 10.3$) were recruited from Cranfield University. 17 participants reported having some experience with robots and automation while 28 reported having no prior experience with robots and automation.

3.1.2 Design

All three studies were an independent design at laboratory conditions. In study 1, participants interacted with a single arm industrial robot to complete an assembly task. In study 2, participants interacted with a twin arm industrial robot to complete an identical task to study 1. In study 3, participants interacted with a single arm industrial robot to complete a pin insertion task.

3.1.3 Materials

Study 1 A single arm industrial robot with a payload capability of 45 kg was used. A laser scanner was used to ensure safe separation between the robot and the participant [6]. For the completion of the assembly task three plastic pipes and three sets of large and small plastic fittings were utilised (Fig. 2).

Study 2 A twin arm industrial robot with a total payload capability of 20 kg was used (Fig. 3). For the completion of the task, only the left-hand side robot gripper was utilised. An identical laser scanner to the previous study was used. For the assembly task, two sets of plastic drain pipes and plastic fittings were utilised identical to study 1.

Study 3 A single arm industrial robot with a payload capability of 200 kg was used. An identical laser scanner
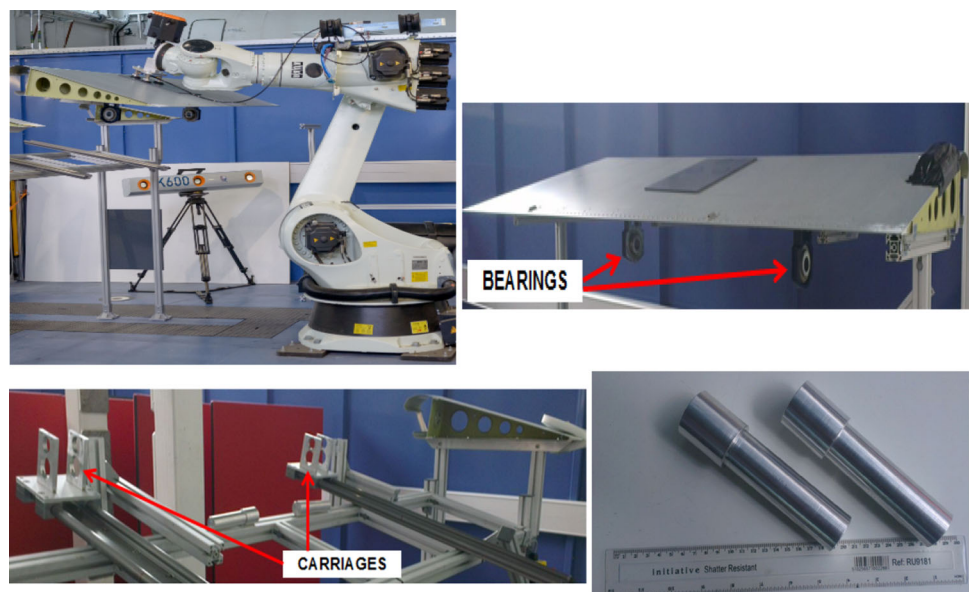
Fig. 2 Materials used for study 1
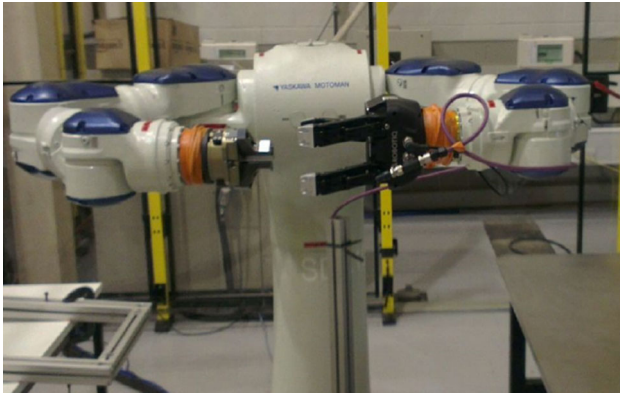
Fig. 3 Industrial robot used for study 2

to the previous study was used. The component lifted by the robot was a representative aerospace sub-assembly. The sub-assembly comprised of two bearings. For securing the sub-assembly a pair of carriages on a stand was designed. For pinning the bearings onto the carriages two identical bearing pins were used (Fig. 4).

3.1.4 Experimental Tasks

Identical tasks were employed for study 1 and 2. The aim was to apply the appropriate plastic fittings on a pipe. The pipes were located next to the robot. The robot picked up one pipe at a time and brought it at participants' standing location. Participants had to attach the plastic fittings on the pipe. Plastic fittings were located next to the participants' standing location. The fittings were disassembled into their respective components in a sequential order. Once both fittings were attached, the completed item was then released by the robot at a drop-off location. Participants completed this task three times in study 1 and two times in study 2.

For study 3 the aim was to secure the sub-assembly's bearings onto the carriages located on the stand utilising two bearing pins. The robot picked up the sub-assembly and positioned it on the stand. Participants walked towards the stand and aligned the carriages, one at a time, to the sub-assembly's bearings by pushing them down. Then participants secured the sub-assembly's bearings on the carriages using the bearing pins. Then participants walked back to their standing

Fig. 4 The industrial robot (*top left*), aerospace sub-assembly (*top right*), carriages (*bottom left*) and the bearing pins (*bottom right*)

position. The robot drove the sub-assembly on the carriages and then released it, indicating the end of the task. Participants completed this task once.

3.1.5 Measures

Data were collected via the 24 item questionnaire developed in the exploratory study. Participants rated each item on a five-point Likert scale and the questionnaire was administered on a computer station. An extract is shown Appendix 4.

3.1.6 Procedure

A standardised procedure was developed identical to all three. Participants were recruited individually from the university campuses. Participants were informed regarding their right to withdraw and anonymity and gave their written consent. Participants were initially taken to a quiet room where they familiarised themselves with the task. Following this participants were taken to the robot cell to interact with the robot. The researcher instructed participants regarding the interaction task. Upon completion, participants completed the 24 item questionnaire. The questionnaire was administered on a computer. Upon completing the questionnaire participants were debriefed and reminded regarding their right to withdraw.

3.2 Analysis of Responses

Analysis proceeded in four steps. Initially, a one way analysis of variance was carried out to identify whether there was a statistical significant difference in the responses obtained between the three studies. Following this, a preliminary reliability analysis was executed to remove any poor items from the analysis. Then a principal component analysis (PCA) was executed to identify the major components. Finally, components were extracted, interpreted and checked for internal consistency.

3.2.1 Exploratory Data Analysis

Collected data were analysed for normal distribution. The Shapiro-Wilk test for trust scores obtained in study 1, $D(60) = 0.979$, $p > 0.05$, study 2, $D(50) = 0.986$, $p > 0.05$; and study 3, $D(45) = 0.969$, $p > 0.05$, indicated no significant difference from normally distributed data. Furthermore, the Levene's statistic for equality of variances indicated no significant difference ($p > 0.05$) suggesting there was no violation of homogeneity of variance. Therefore parametric analysis was used. Table 2 presents the descriptive statistics of the three groups.

Table 2 Descriptive statistics of the three groups

Descriptive Statistics				
Groups	N	Mean	SD	SE mean
Study 1	60	96.75	8.9	1.2
Study 2	50	93.88	9.6	1.4
Study 3	45	95.51	10.2	1.5

On average, participants in study 1 experienced higher trust in the robotic teammate (M = 96.75, SE = 1.160), when compared to the participants in study 2 (M = 93.88, SE = 1.359) and study 3 (M = 95.51, SE = 1.527). However, this difference was not statistically significant $F(2) = 1.228$, $p > 0.05$. Therefore, the data were merged into a single dataset, providing 155 cases for further analysis.

3.2.2 Preliminary Reliability Analysis

In order to improve the reliability of the questionnaire, a preliminary reliability analysis was carried out. Prior to analysis the data were transformed (as appropriate) so that in all cases a high score represented higher trust. Initial reliability analysis generated a Cronbach alpha of 0.811 which is well within the generally acceptable level of 0.7 suggested in the literature [31]. The next step taken to improve the scale is to identify the items that do not contribute to the overall reliability. No removal of any item would increase Cronbach alpha by a significant amount. The decision to remove items was made on the basis of the 'Corrected item-total correlation'. This is a correlation between the item score and the overall test score, excluding the item in question from the total score. This correction is performed to avoid inflation of the item-total correlation [32]. Lowenthal [33] suggests a removal threshold of between 0.15 and 0.30. However, because of the exploratory nature of this questionnaire the mean item-total correlation was taken as an indicator. The mean item-total correlation is 0.374 giving a higher cut-off margin than the one suggested in literature. Applying this rule resulted in the removal of eleven items. A reliability analysis of the remaining 13 items was carried out generating a Cronbach alpha of 0.84 indicating increased reliability of the scale. These 13 items were then subjected to a PCA.

3.2.3 Principal Component Analysis

The data were then subjected to PCA using SPSS (version 20). Data were subjected to a Varimax rotation to produce a more interpretable solution [34]. Kaiser's criterion (components with an Eigenvalue in excess of one) was used to determine the maximum number of components extracted. For an item to be deemed to load onto one of the extracted

components it was required to have a loading in excess of 0.45 [35]. Items loading onto no components or onto two (or more) components were excluded from the final solution. Also, items with low communalities were excluded from the solution [35]. After removing an item PCA was re-run to identify the new factor structure. From the PCA three components were extracted accounting for 63.5 % of the total variance in the sample. A Kaiser-Meyer-Olkin (KMO) statistic of 0.847 was achieved which is above Kaiser's minimum cut-off level of 0.5 indicating sample size is sufficient [36]. Also, Bartlett's test of sphericity is statistically significant ($\chi2(45) = 465.6$, $p < 0.001$). This result indicates that there is significant correlation within the dataset so components are unlikely to occur by chance. The final component structure is shown in appendix 5. Three major components were generated and items loaded clearly on each of the components. Therefore, no further item removal was necessary. The next step was to interpret the components identified and investigate their internal consistency (reliability).

3.2.4 Component Interpretation and Reliability Analysis

Three components emerged from the PCA. Component 1, was termed 'Safe co-operation', consisted of four items and had a Cronbach's alpha of 0.802. Component 2 was termed 'Robot and gripper reliability', consisted of four items and had an alpha value of 0.712. Component 3 was termed 'Robot's motion and pick-up speed', consisted of two items and resulted in a Cronbach's alpha of 0.612.

3.2.5 Summary of Results

The statistical analysis enabled the development of a ten item psychometric scale to measure the development of trust in industrial HRC. In summary, trust in industrial HRC is, primarily, affected by three key factors (components), each of

which is assessed with a number of items. The developed trust scale is summarised in Table 3.

The following section discusses the output of the psychometric scale and presents the practical implications, as well as, a user's guide for practitioners.

4 Discussion and Practical Implications

The output of this work provides a number of theoretical and practical implications. These are discussed in Sects. 4.1 and 4.2 respectively.

4.1 Discussion on the Theoretical Contributions of the Scale

The statistical analysis suggests that trust in industrial HRC depends on three components: safe co-operation, robot and gripper reliability and robot's motion and pick-up speed. The components exhibited fairly good internal consistency. Components 1 and 2 are within the general acceptable cut-off limit of 0.7 suggested in the literature [31] indicating good reliability. Although component 3 exhibited an alpha value (0.612) lower than the minimum acceptable limit, Kline [37] suggests that for psychological constructs values lower than 0.7 can also be accepted. At the same time, this alpha value is acceptable for newly developed scales [38] particularly given the small number of items in this component (two).

One of the major components identified through the analysis was safety during the co-operation between the human and the industrial robot. This finding is consistent with earlier work, suggesting that a positive level of perceived safety can be a key element for the successful introduction of robots in human environments [39,46]. A recent study by Shiomi, Zanlungo, Havashi and Kanda [46] highlighted that if a robot is to be successfully integrated within the human environ-

Table 3 The developed psychometric scale to measure trust in industrial HRC

Scale item	Major components
The way the robot moved made me uncomfortable	Robot's motion and pick-up speed
The speed at which the gripper picked up and released the components made me uneasy	
I trusted that the robot was safe to cooperate with	Safe Co-operation
I was comfortable the robot would not hurt me	
The size of the robot did not intimidate me	
I felt safe interacting with the robot	
I knew the gripper would not drop the components	Robot and gripper reliability
The robot gripper did not look reliable	
The gripper seemed like it could be trusted	
I felt I could rely on the robot to do what it was supposed to do	

ment, it must be first perceived as safe by the human partner. In this work, the items grouped in this component indicate that both mental (impact of the robot size) and physical safety (not being injured by the robot) is important during a HRC task in industry which is in line with previous literature [40]. This is particularly important for the industrial context where human operators will be required to work in close proximity with industrial robots. In some occasions, such as the one used for study 3, these robots can have a very high payload capability and their size can be intimidating. It appears that ensuring operators are exposed to a collaborative scenario where safety is facilitated, can generate a positive feeling of safety. This in turn can assist the human operator to develop trust in the robotic partner.

The performance aspects of the robotic system and specifically, the reliability of the robot and the gripper was the second trust related component. Robot reliability is in line with previous and more recent literature [15,47]. In a meta-analysis by Hancock and colleagues [24] robot performance factors (e.g. reliability) had the highest impact on trust. Furthermore, van der Brule and colleagues [47] reconfirmed that a robot's task performance influences human trust. The findings of this study highlight once again the criticality of a reliable robot system. An unreliable robot will eventually decrease operator's trust which in turn will be detrimental for accepting and using the robot. Also, considering that humans are far more sensitive to automation errors thus leading to a significant drop in trust [27], robot reliability becomes a very important aspect.

Interestingly, the reliability of the gripping mechanism appeared to have an impact on trust. To our knowledge, this context specific aspect has not appeared in previous literature. This is of particular relevance to industrial HRC, since the gripper is a vital component of an industrial robot. The gripping mechanism is the mean with which the robot will manipulate components and interact with the human partner in a collaborative task. As industrial robots come in a variety of gripping mechanisms depending on the task being utilised for, findings suggest that the reliability of the gripping mechanism is an important determinant for trust development. When the reliability of the gripping mechanism decreases, human trust in the robotic partner decreases.

The third trust component was relevant to the robot's motion and the component pick-up speed. It appears that the motion of the robot is an important factor for the development of trust. This is in line with previous research indicating that robot's movement can assist the human partner to predict and anticipate robot's intentions [41,42]. A fluent, non-disruptive robot movement can put the human partner at ease and foster trust. This is particularly important for an industrial environment where the robot will be collaborating in close proximity with a human operator. Furthermore, industrial settings can

be cluttered with other operators therefore it is important for other operators to predict the robot's movement. Also, the final component suggested that the speed at which the gripping mechanism picks-up components has an impact on the development of trust. Similar with the previous component (robot and gripper reliability) the robot's gripping mechanism appears to have an important role in the development of trust.

In addition, the statistical analysis indicated that that the appearance of the robot did not emerge as a contributing component to trust development. Previous literature in the domain of social robotics provides contradicting results in terms of the effects of robot appearance on user preferences; some suggest robots should not be too human-like in appearance whereas others indicate that more human-like appearance can engage people more [43–45,48]. Astrid and colleagues [48], for example, investigated participants' evaluations of very human-like robots as well as their attitudes towards these robots. Their results showed both positive and negative attitudes towards very human-like robots. Similarly, Prakash and Rogers [49] found that human perception of a robot tends to vary based on the robot's human-likeness. According to their findings, humans tend to over-generalise the capabilities of a very human-like robot. Further on this, earlier literature stressed that anthropomorphic appearance should be treated with care in order to match the appearance of the robot with its abilities without generating unrealistic expectations to the human user [39]. This finding possibly indicates that people perceive industrial robots as tools used to complete a task. Therefore it appears that robot appearance for industrial HRC is not a major contributor to trust development when compared to social robots used as social companions.

4.2 Discussion on the Practical Implications and a User Guide for Practitioners

The output of this work has significant practical implications. First, to our knowledge, this is the only empirically developed psychometric scale for measuring trust in industrial HRC.

Second, this scale can be a powerful tool for system designers and organisations aiming to implement industrial HRC. It provides guidance on how system characteristics can affect operators' perception of trust. For instance, the scale identified three key system aspects fostering trust industrial HRC: safety, robot and gripper reliability and robot's motion and gripper pick-up speed. These three areas appear to be the major determinants for trust development in an industrial HRC scenario.

Third, the scale can be used to identify the relationship of each individual operator and raise awareness regarding personal tendencies. For example, poor scores on robot and gripper reliability might identify those operators in need

for further training regarding the capabilities and technical aspects of the gripping mechanism. For this purpose, a user guide has been created to assist practitioners in using this psychometric scale to: (i) administer the trust scale post task and collect trust results, (ii) analyse the collected scores and (iii) interpret the scale output and take appropriate action accordingly. This is presented in Appendix 6. The user guide is segregated in four parts:

- Part A: Participant instructions: Provides instructions to participants on how to complete the questionnaire.
- Part B: Participant demographic: Provides a short demographic section.
- Part C: Questionnaire: The developed scale items were randomly placed in a questionnaire (five point scale) and it can be immediately administered.
- Part D: Instructions for the assessor: This section provides a five step process to enable the assessor to correctly analyse the results, interpret the output and take appropriate actions.

5 Future Work

The results of this study can provide the basis for further work on trust development in industrial HRC. First, because this study was making a first attempt to understand trust development in industrial HRC, university students and staff participants took part rather than factory workers. The majority of the population did not have any prior experience with industrial robots or automation before. Therefore, it is important to validate the results using individuals who have an in-depth understanding of industrial robots or manufacturing automation. Second, this study was carried out in laboratory conditions. Future work should be geared to investigate whether the results and trends from this study apply in a real-world scenario. Third, future research could investigate each of the components identified in the scale individually. For instance, scale identified robot and gripper reliability as important determinants of trust. Further research could investigate the impact on human trust under varying levels of gripper reliability. This could provide a trust region for which collaboration is optimised.

Appendix

Appendix 1: Interview Template

Section	Main question	Probe
Introduction	Can you talk to me about your first thoughts regarding the interaction with this robot?	Why did you feel this way?
Robot related	Did you feel you could rely on the robot to hand you over the components safely?	Why?
	Can you talk to me about the robot's ability to hand you the components?	Why? Can you tell me more?
	How did the appearance of the robot influence your trust?	Why?
Safety	Did you have any concerns when you interacted with the robot?	What? Why?
Other topics	Considering the task you have just completed, what has encouraged you to trust the robot?	Can you talk to me more about this?
	Is there anything else about the robot that encouraged you to trust this robot?	Why? Can you tell me more?

Appendix 2: Coding Template

Element	Trust-related theme (code sign)	Lower level theme (code sign)
Robot (R)	Robot's performance (R1)	Robot motion (R1m)
		Robot and gripper reliability (R1r)
	Robot's physical attributes (R2)	Robot size (R2s)
		Robot appearance (R2a)
Human (H)	Safety (H1)	Personal safety (H1p)
		Safe programming of the robot (H1prog)
	Experience (H2)	Prior interaction experiences (H2int)
		Robot mental models (H2mm)
External (E)	Task (E1)	Complexity of the task (E1comp)

Appendix 3: Item Generation

Items	Direction
The way the robot moved made me uncomfortable	−
I was not concerned because the robot moved in an expected way	+
The speed of the robot made me uncomfortable	−
The speed at which the gripper picked up and released the components made me uneasy	−
I felt I could rely on the robot to do what it was supposed to do	+
I knew the gripper would not drop the components	+
The design of the robot was friendly	+
I believe the robot could do a wider number of tasks than what was demonstrated	+
I felt the robot was working at full capacity	−
The robot gripper did not look reliable	−
The gripper seemed like it could be trusted	+
The size of the robot did not intimidate me	+
I felt safe interacting with the robot	+
I was comfortable the robot would not hurt me	+
I trusted that the robot was safe to cooperate with	+
I had faith that the robot had been programmed correctly	+
The way robots are presented in the media had a negative influence on my feelings about interacting with this robot	−
I had no prior expectations of what the robot would look like	+
I don't think any prior experiences with robots would affect the way I interacted with the robot	+
If I had more experiences with other robots I would feel less concerned	-
I was uncomfortable working with the robot due to the complexity level of the task	-
If the task was more complicated I might have felt more concerned	-
I might not have been able to work with the robot had the task been more complex	-
The task made it easy to interact with this robot	+

Appendix 4: Extract of the Rating Scale

PLEASE RATE THE FOLLOWING STATEMENTS BASED ON THE TASK YOU HAVE JUST COMPLETED

Q1	The way the robot moved made me uncomfortable				
	1	2	3	4	5
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Q2	The task made it easy to interact with this robot				
	1	2	3	4	5
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Q3	I was not concerned because the robot moved in an expected way				
	1	2	3	4	5
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Appendix 5: Trust Scale Components

		Component		
		1	2	3
1	The way the robot moved made me uncomfortable			0.759
2	The speed at which the gripper picked up and released the components made me uneasy			0.848
3	I knew the gripper would not drop the components		0.651	
4	The robot gripper did not look reliable		−0.828	
5	I trusted that the robot was safe to cooperate with	0.688		
6	The gripper seemed like it could be trusted		0.793	
7	I was comfortable the robot would not hurt me	0.782		
8	The size of the robot did not intimidate me	0.754		
9	I felt safe interacting with the robot	0.787		
10	I felt I could rely on the robot to do what it was supposed to do		0.506	
	Reliability analysis output: cronbach alpha value	0.802	0.712	0.612

Appendix 6: User Guide to the Trust Scale Questionnaire

QUESTIONNAIRE

PART A: PARTICIPANT INSTRUCTIONS

The aim of this questionnaire is to assess the effectiveness of the collaboration between yourself and the industrial robot you are collaborating with.

The questionnaire requires you to rate 10 statements based on **a task you accomplish in collaboration with an industrial robot (take the most frequent task you work on)**. The statements are in bold and need to be rated on a five-point scale ranging from: *Strongly Disagree* to *Strongly Agree*. Read each statement and then check (X) the preferred box. An example is given below:

	I drive carefully even in roads I do not know well				
i				X	
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Please remember, it is the effectiveness of the collaboration being assessed **and not you**.

For anonymity purposes a participant number has already been provided at the top of the next page. This is to stress that your individual responses will be anonymised and personal details will not be shared. All data will be stored and maintained by Cranfield in accordance with the Data Protection Act (1998).

A short participant demographic form is provided below.

PART B: PARTICIPANT DEMOGRAPHIC

PARTICIPANT ID: ____Y____   Age: _____   Sex: _____

Please describe the collaborative industrial robot used (e.g. twin arm medium payload, single arm large payload, three finger clamp single arm etc):

_____

_____

_____

PART C: QUESTIONNAIRE   **PARTICIPANT ID: [Give a number]**

A	The way the robot moved made me uncomfortable				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

B	I felt I could rely on the robot to do what it was supposed to do				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

C	The speed at which the gripper picked up and released the components made me uneasy				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

D	I felt safe interacting with the robot				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

E	**I knew the gripper would not drop the components**				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

F	**The size of the robot did not intimidate me**				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

G	**The robot gripper did not look reliable**				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

H	**I was comfortable the robot would not hurt me**				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

I	**I trusted that the robot was safe to cooperate with**				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

J	**The gripper seemed like it could be trusted**				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Part D: Instructions for the Assessor

The statements in this survey have been randomised to reduce participant bias. Follow these five steps post administration to analyse the output.

Step 1: Group the Statements Together into their Major Components as shown below:

Questionnaire number	Major components
A	Robot's motion and pick-up speed
C	
D	Safe Co-operation
F	
H	
I	
B	Robot and gripper reliability
E	
G	
J	

Step 2: Score the Individual Statements:

*The following scoring scheme assumes that the individuals were exposed to an industrial robot of 100 % reliability (e.g. no failures, no abnormal motion)

Scale points	Use this scoring scheme for statements: B, D, E, F, H, I, J Score	Use this scoring scheme for statements: A, C, G Score
Strongly agree	5	1
Agree	4	2
Neutral	3	3
Disagree	2	4
Strongly disagree	1	5

Step 3: Deriving the Trust Score for Each Component and the Total Trust Score

ID	Trust component	Individual score	Minimum score possible	Maximum score possible
1	Perceived Robot's motion and pick-up speed	X	2	10
2	Perceived Safe co-operation	Y	4	20
3	Perceived Robot and gripper reliability	Z	4	20
Totall trust score		X+Y+Z	10	50
Date undertaken		DD/MM/YYYY		

Step 4: Interpreting the Results

First, observe the total trust score. If the total trust score is low (e.g. less than 25), then this could be an indication the individual does not trust the robot to collaborate with. This could have severe implications on operational effectiveness and efficiency. At the same time, if the total trust score is very high (e.g. close to 50), then this could potentially indicate the individual is overly relying on the robot which could lead to complacency.

Secondly, observe the scores achieved for each trust component separately. Identify any poorly rated trust components. For example, a particular worker might have scored high on component 3 ('Perceived robot and gripper reliability') but poorly on component 2 ('Perceived safe Co-operation') and 1 ('Perceived robot's motion and pick-up speed'). This will enable to identify the specific areas requiring actions, e.g. refresher training course.

Step 5: Identifying Emerging Trends

To identify emerging trends and patterns it is suggested to:

- Log results for each individual
- Administer the questionnaire every four months. Consider re-randomising the order to the statements on the questionnaire to avoid participant bias
- Compare results between each quarter
- Communicate to the individual the results of the questionnaire. If the individual is identified to be in the 'low' or 'very high' region, follow up with an informal conversation to understand the reasons behind this. This will enable to take mutually agreed actions.

*Data for Comparison

The table below presents the average score and standard deviation of each of the questionnaire items is provided following the experimental trials at Cranfield and Loughborough Universities (Sample size: 155). These data can be used as an early comparison with the collected data.

Questionnaire item	Average score	SD
A	1.8	0.8
C	1.8	0.9
D	4.3	0.7
F	4.1	0.9
H	4.3	0.8
I	4.3	0.7
B	4.1	0.9
E	3.7	1.1
G	1.9	0.9
J	4	0.8

References

1. Ding Z, Hon B (2013) Constraints analysis and evaluation of manual assembly. CIRP Ann Manuf Technol 62(1):1–4
2. Hägele M, Schaaf W, Helms E (2002) Robot assistants at manual workplaces: Effective co-operation and safety aspects. *International symposium on robotics ISR 2002 / CD-ROM*. October 7–11, Stockholm
3. Schraft RD, Meyer C, Parlitz C, Helms E (2005) Powermate—a safe and intuitive robot assistant for handling and assembly tasks. *IEEE international conference on robotics and automation*, IEEE, Barcelona, pp 4047–4079
4. Santis AD, Siciliano B, Luca AD, Bicchi A (2008) Atlas of physical human-robot interaction. Mech Mach Theor 43(3):253–270
5. Unhelkar VV, Siu HC, Shah JA (2014) Comparative performance on human and mobile robotic assistants in collaborative fetch-and-deliver tasks. Human Robot interaction 2014. ACM, Bielefeld
6. ISO (2011) Robots and robotic devices-safety requirements for industrial robots, Part 1: robots. international standards organisation, Geneva
7. Bortot D, Born M, Bengler K (2013) Directly or detours? How should industrial robots approach humans?. ACM/IEEE international conference on human-robot interaction (HRI 2013), IEEE, Tokyo
8. Walton M, Webb P, Poad M (2011) Applying a concept for robot-human cooperation to aerospace equipping processes, *SAE International*
9. Parasuraman R, Riley V (1997) Humans and automation: use, misuse, disuse, abuse. Hum Fact 39(2):230–253
10. Freedy A, de Visser E, Weltman G, Coeyman N (2007) Measurement of trust in human-robot collaboration. In: *Proceedings of the 2007 international conference on collaborative technologies and systems*, Orland
11. Chen JY, Barnes MJ (2014) Human-agent teaming for multirobot control: a review of human factors issues. IEEE Trans Hum Mach Syst 44(1):13–29 IEEE
12. Groom V, Nass C (2007) Can robots be teammates? Benchmarks in human–robot teams. Interact Stud 8(3):483–500
13. Park E, Jenkins Q, Jiang X (2008) Measuring trust of human pperators in new generation rescue robots. *Proceedings of the 7th JFPS international symposium on fluid power*, Toyom, pp 15–18
14. de Visser EJ, Parasuraman R, Freedy A, Freedy E, Weltman G (2006) A comprehensive methodology for assessing human-robot team Performance for use in training and simulation. Proceedings of the 50th Human factors ergonomics society, San Francisco, pp 2639–2643
15. Lee JD, See KA (2004) Trust in automation: designing for appropriate reliance. Hum Fact 46(1):50–80
16. Verberne F, Ham J, Midden C (2012) Trust in smart systems: sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars. Hum Fact 54(5):799–810
17. Mazney D, Reichenbach J, Onnasch L (2012) Human performance consequences of automated decision aids: the impact of degree of automation and system experience. J Cogn Eng Decis Mak 6(1):57–87
18. Parasuraman R, Molloy R, Singh I (1993) Performance consequences of automation-induced 'complacency'. Int J Aviat Psychol 3(1):1–23
19. Dzindolet MT, Pierce LG, Beck HP, Dawe LA, Anderson WB (2001b) Predicting misuse and disuse of combat identification systems. Mil Psychol 13(3):147–164
20. de Visser E, Krueger F, McKnight P, Scheid S, Smith M, Chalk S et al. (2012) The world is not enough: trust in cognitive agents. In: Proceedings of the 56th annual HFES meeting, pp 263–267
21. Chen JY, Barnes MJ (2012) Supervisory control of multiple robots: effects of imperfect automation and individual differences. Hum Fact 54(2):157–174
22. Desai M, Stubbs K, Steinfeld A, Yanco H (2009) Creating trustworthy robots: lessons and inspirations from automated systems. Paper presented at The AISB convention: new frontiers in human-robot interaction. Edinburgh
23. Yagoda RE, Gillan DJ (2012) You want me to trust a ROBOT? The development of a human-robot interaction trust scale. Int J Soc Robot 4(3):235–248
24. Hancock PA, Billings DR, Oleson KE, Chen JY, De Visser E, Parasuraman R (2011) A meta-analysis of factors influencing the development of human-robot trust. Aberdeen proving ground, MD 21005-5425: US Army Research Laboratory
25. Singh IL, Molloy R, Parasuraman R (1993) Automation-induced "complacency": development of a complacency-potential rating scale. Int J Aviat Psychol 3(2):111–122
26. Muir BM, Moray N (1996) Trust in automation. Part 1. Experimental studies of trust and human intervention in a process control simulation. Ergonomics 39(3):429–460
27. Jian X, Bisantz AM, Drury CG (2000) Foundations for an empirically determined scale of trust in automated systems. Int J Cognit Ergon 4(1):53–71
28. Master R, Gramopadhye AK, Melloy BJ, Bingham J, Jiang X (2000) A questionnaire for measuring trust in hybrid inspection systems. Paper presented at the industrial engineering research conference. Dallas
29. Schaefer KE (2013) *The perception and measurement of human-robot trust*. Doctoral dissertation, University of Central Florida Orlando, Florida
30. King N (1998) Template analysis. In: Gillian S, Catherine C (eds) Qualitative methods and analysis in organizational research: a practical guide. Sage Publications Ltd, Thousand Oaks, pp 118–134
31. Nunnally JC (1978) Psychometric theory, 2nd edn. McGraw-Hill, New York
32. Kline TJB (2005) Psychological testing: a practical approach to design and testing. Sage Publications Inc, London
33. Lowenthal KM (1996) An introduction to psychological tests and scales. UCL Press, London
34. Harris D, Chan-Pensley J, McGarry S (2005) The development of a multidimensional scale to evaluate motor vehicle dynamic qualities. Ergonomics 48(8):964–982
35. Hair JF, Anderson RE, Tatham RL, Black WC (1998) Multivariate data analysis. Prentice Hall, Upper Saddle River
36. Kaiser HF (1974) An index of factorial simplicity. Psychometrika 39(1):31–36
37. Kline P (1999) The handbook of psychological testing, 2nd edn. Routledge, London
38. deVellis RF (1991) Scale development: theory and application. Sage Publishing, Newbury Park
39. Bartneck C, Kulic D, Croft E (2009) Measuring the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int J Soc Robot 1:71–81
40. Inoue K, Nonaka S, Ujiie Y, Takubo T, Arai T (2005) Comparison of human psychology for real and virtual mobile manipulators. *IEEE international workshop on robot and human interactive communication, ROMAN 2005*, IEEE, pp 73–78
41. Huber M, Rickert M, Knoll A, Brandt T, Glasauer S (2008) Human-robot interaction in handing-over tasks. 17th IEEE international symposium on robot and human interactive communication, pp 107–112
42. Mayer MP, Kuz S, Schlick CM (2013) Using anthropomorphism to improve the human-machine interaction in industrial environments (part II). In: Duffy VG (ed) Digital human modeling and applications in health, safety, ergonomics, and risk management. Human body modelling and ergonomics. Springer, Berlin, pp 93–100

43. Broadbent E, Stafford R, MacDonald B (2009) Acceptance of healthcare robots for the older population: review and future directions. Int J Soc Robot 1(4):319–330

44. Bartneck C, Kanda T, Mubin O, Mahmud AA (2009) Does the design of a robot influence its animacy and perceived intelligence? Int J Soc Robot 1(2):195–204

45. Li D, Rau PL, Li Y (2010) A cross-cultural study: effect of robot appearance and task. Int J Soc Robot 2(2):175–186

46. Shiomi M, Zanlungo F, Hayashi K, Kanda T (2014) Towards a socially acceptable collision avoidance for a mobile robot navigating among pedestrians using a pedestrian model. Int J Soc Robot 6(3):443–455

47. van den Brule R, Dotsch R, Bijlstra G, Wigboldus DH, Haselager P (2014) Do robot performance and behavioral style affect human trust? Int J Soc Robot 6(4):519–531

48. Rosenthal-von der Pütten AM et al, Krämer NC (2015) Individuals' evaluations of and attitudes towards potentially uncanny robots. Int J Soc Robot 7(5):1–26

49. Prakash A, Rogers WA (2014) Why some humanoid faces are perceived more positively than others: effects of human-likeness and task. Int J Soc Robot 7(2):309–331

George Charalambous is currently a human factors consultant for Interfleet Technology (Member of the SNC Lavalin Group). George has recently completed a PhD at Cranfield University funded by the EPSRC Centre for Innovative Manufacturing in Intelligent Automation. George's research interests revolve around the exploration of the human factors to enable the successful implementation of human-robot/automation collaboration in safety critical systems, such as rail automated systems.

Sarah Fletcher leads the Industrial Psychology and Human Factors group in the Advanced Systems Centre at Cranfield University as well as the human centred research of the EPSRC Centre for Innovative Manufacturing in Intelligent Automation. She has over 15 years experience of conducting psychological and human factors research particularly in complex systems and manufacturing contexts.

Philip Webb is a Royal Academy of Engineering and Airbus Professor and is the head of the Advanced Systems Centre at Cranfield University. He is also Deputy Director of the EPSRC Centre for Innovative Manufacturing in Intelligent Automation. He has over 20 years of experience in research into advanced automation and robotics. His particular specialism is the application of automation in the aerospace industry and he has worked with most of the major aerospace companies including Airbus, BAE System and Rolls-Royce.