HR-Recycler: Hybrid Human-Robot RECYcling plant for electriCal and eLEctRonic equipment

D7.1 – Human factors modelling in human-robot interaction (version 1)

WP number and title	WP7 – D7.1 – Human factors modelling in human-robot interaction (version 1)
Lead Beneficiary	TEC
Contributor(s)	TEC, IBEC, CERTH
Deliverable type	Report
Planned delivery date	30/11/2019
Last Update	31/03/2020
Dissemination level	PU

Disclaimer

The HR-Recycler Consortium consists of the following partners:

Participant No	Participant organisation name	Short Name	Type	Country
1	Centre for Research and Technology Hellas CERTH - ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS	CERTH	RTO	GR
2	FUNDACIO INSTITUT DE BIOENGINYERIA DE CATALUNYA	IBEC	RTO	ES
3	TECHNISCHE UNIVERSITAET MUENCHEN	TUM	RTO	DE
4	COMAU SPA	COMAU	IND	IT
5	FUNDACION TECNALIA RESEARCH & INNOVATION	TEC	RTO	ES
6	ROBOTNIK AUTOMATION SLL	ROB	SME	ES
7	FUNDACION GAIKER	GAIKER	RTO	ES
8	SADAKO TECHNOLOGIES SL	SDK	SME	ES
9	DIGINEXT	DXT	IND	BE
10	VRIJE UNIVERSITEIT BRUSSEL	VUB	RTO	BE
11	INDUMETAL RECYCLING, S.A.	IND	SME	ES
12	INTERECYCLING - SOCIEDADE DE RECICLAGEM SA	INT	SME	PT
13	BIANATT ANAKYKLOSI AIIE ANONIMI BIOMICHANIKI EMPORIKI ETAIRIA	BNTT	IND	GR

Document History

VERSION	DATE	STATUS	AUTHORS, REVIEWER	DESCRIPTION
0.1	16/04/2019	Draft	Sara Sillaurren (TEC)	Table of Contents
0.2	25/09/2019	Draft	Sara Sillaurren, Leire Bastida, Erlantz Loizaga (TEC)	Contributions from TECNALIA
0.3	08/10/2019	Draft	Sara Sillaurren, Leire Bastida, Erlantz Loizaga (TEC)	Updated table of content and contributions from TECNALIA
0.4	10/10/2019	Draft	Ioannis Chatzikonstantinou (CERTH)	Initial draft of section 2.2
0.5	11/10/2019	Draft	Sara Sillaurren (TEC)	Updated contents by TEC
0.6	29/10/2019	Draft	Sara Sillaurren (TEC)	Version ready to IBEC contribution
0.7	11/11/2019	Draft	Sara Sillaurren, Leire Bastida, Erlantz Loizaga (TEC) Vicky Vouloutsi (IBEC)	Contribution by IBEC and SotA Trust
0.8	13/11/2019	Draft	Sara Sillaurren, Erlantz Loizaga (TEC)	Updated section 4.1 by TEC
0.9	25/11/2019	Draft	Vicky Voulotsi (IBEC), Dirk Wollherr (TUM), Sara Sillaurren, Leire Bastida and Erlantz Loizaga (TEC)	Integrated comments by TUM and new section 5 from IBEC
1.0	28/11/2019	Final	Sara Sillaurren, Leire Bastida (TEC)	Final review of document
1.1	09/03/2020	Reviewed	Sara Sillaurren (TEC)	Modifications after 1st review comments from PO
1.2	17/03/2020	Reviewed	Ioannis Chatzikonstantinou (CERTH)	Modifications after 1st review comments from PO
1.3	30/03/2020	Reviewed	Vicky Vouloutsi (IBEC)	Modifications after 1st review comments from PO

Definitions, Acronyms and Abbreviations

ACRONYMS / ABBREVIATIONS	DESCRIPTION
AGV	Automated Guided Vehicle
ANS	Autonomic Nervous Systems
AR	Augmented Reality
BVP	Blood Volume Pulse
ECG	Electrocardiography
EEG	Electroencephalogram
EM	Emotional Model
EMG	Electromyography
ERP	Event-related potential
FPD	Flat Panel Display
GPS	Global Positioning System
GSR	Galvanic Skin Response
HF	Human Factors
HHI	Human-Human Interaction
HRC	Human-Robot Collaboration
HRI	Human-Robot Interaction
HRV	Heart Rate Variability
IMU	Inertial Measurement Unit
KPI	Key Performance Indicator
LPT	Line Print Terminal
PC	Personal Computer
PPG	Photoplethysmography
RESP	Respiration
RR intervals	R is a point corresponding to the peak of the QRS complex of the ECG wave and RR is the interval between successive Rs
SotA	State-of-the-Art
TEMP	Temperature
UC	Use Cases
UX	User eXperience

WEEE	Waste Electric and Electronic Equipment

Table of Contents

List of Figures

List of Tables

Executive Summary

The deliverable 7.1 "Human factors modelling in human-robot interaction (version 1)" is a deliverable of Work Package 7 "Human-Robot Collaboration schemes", more concretely, from the Task 7.1 "Human factors for UX analysis". The objective of the task is to provide explicit models of the human factors by integrating human cognition and behaviour variables from physical, emotional and cognitive perspectives in order to enable the human-robot interaction to react to context and human restrictions individually.

For achieving these objectives, the following subtasks have been completed:

- Definition of requirements for an effective communication between robots and human co-workers, including the link between communication and collaboration and how the first can influence the latter.
- Description of the dynamics in the interaction between humans and robots based on the findings and definitions in WP3.
- Initial specification of the Human Factors Worker model based on the Trust factor. This model will be connected and included in the general Worker model from T4.2.
- Organization of a set of pilot studies with end-users to measure Trust factor.
- Initial assessment of robot social behaviours.

State of the Art: the state of the art for human factors modelling in the human-robot interaction is summarized below (see sections 2 and 3.1 for more details):

- Worker Model is a *work-related abstract model* that refers to a high-level description of several worker characteristics. Several projects have defined general workers models, as for example, Factory2fit European project[1].
- The *introduction of Human Factors* (emotions and cognitive processes) issues in Worker Model[2] is a relatively new subject and it's related in effects the system has on the human[3] or the human interacting with robots, in terms of, for example fatigue, stress…
- In industrial settings, adaptive solutions have been proposed to ensure flexibility, safe operation, robot programming, as well as optimization in task planning and execution. The incorporation of social cognition (where human factors are included) has been added in industrial collaborative robots where the robot's behavior has been evaluated in terms of trust, comfort and robot performance and team fluency [86]-[88]. Several projects have incorporated adaptive behaviors to robots, as can be seen in the Rossini EU project[4].

---

[1] Factory2Fit EU Project, D2.1 - Empowering and Participatory Adaptation of Factory Automation to Fit for Workers, https://factory2fit.eu/wp-content/uploads/2017/10/Factory2Fit-D2.1-Dynamic-Worker-Model.pdf

[2] M. Iza, 2019, "Emotional Human-Computer Interface:Are emotional inferences forward?", AIC 2019 https://pdfs.semanticscholar.org/2401/42c24dd4fc74d618f1d7d87591ef503dac0a.pdf

[3] R. McLeod, 2004, Human factors assessment model validation study, Health and Safety Executive 2004 Report https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=2ahUKEwiG24HKgeDnAhWD3OAKHe8cCAAQFjABegQIBBAB&url=https%3A%2F%2Fwww.hse.gov.uk%2Fresearch%2Frrpdf%2Frr194.pdf&usg=AOvVaw2RJCTYbSt--JI1MpSarPxS

[4] Rossini (RObot enhanced SenSing, INtelligence and actuation to Improve productivity and job quality in manufacturing) EU project, D2.1 ROSSINI State of the Art Analysis

- *Interaction using Augmented Reality* has found application in industrial settings as means of increasing process awareness and enhancing safety.
- Optimal robotic actions planning is crucial for HRC. Action plans are generated based on the robot's perceptions, capabilities, task goals [95], cost [94] and allow adaptation to changes in the work-flow [96].
- *Robot/AGVs navigation in industrial environments* is a great challenge since typically dedicated zones are foreseen in order to avoid collisions with existing resources (other AGVs) and humans [5].
- AGVs navigation in factory environments can lead to *delays in intra-factory transportation* and thus selected solutions should follow global trajectory optimization and not just obstacle avoidance [6].
- AR-based communication is considered a *viable alternative in industrial environments* where implicit communication channels are challenging to employ.
- *AR-based HRC* has been explored in industrial applications in order to enhance productivity and safety [16]-[18].
- *Trust is a major factor to determine effective collaboration* among agents, both human-to-human and human-to-machine [36]-[45]
- Only factors within the *situational or learned trust dimensions* are considered with a demonstrated correlation with trust [49].
- Typical industrial robot programming includes instructive systems and learning by demonstration [89], [90]. Recent approaches to robot learning include active learning. In this paradigm, to improve learning rates in uncertain situation, systems are endowed with agency, that is the ability to request certain information and how to acquire it [91]-[93]. However, such approach is not yet common in industrial robots.

HR-Recycler contribution: the contribution of this deliverable beyond the state of the art is a novel Trust Factor Worker model (outlined in sections 3.2 and 4):

- A new *Human Factors Worker model* will be developed for theoretically asset the human operator's state based on psychophysiological measures (it will be reported in section 3.2)
- This HF Worker model will be closely related with the one developed inside WP4 (in T4.2)
- To achieve successful interactions and collaboration as well as ensure user predictability and perceived safety, HR-Recycler will incorporate *social cognition and adaptive behaviors* to robots in a centralized system. More specifically, HR-Recycler will develop a novel interaction engine that will employ behaviors that ensure effective communication and are adapted based on the social affordances extracted by the Worker Model and the user's state.
- HR-Recycler will develop advanced AR-based HRC that will go beyond one-to-one human-robot interaction, offering *extended multi-workstation process overview capabilities* that will enhance cognitive abilities of workers to better supervise ongoing processes. In addition, projective AR will augment factory floor safety by providing robot and AGV intention and warning indications, through projective AR techniques. Projective AR will have a positive effect in human comfort as additional wearable equipment is not necessary for perception of immediate alerts. Projective alerts will be developed to prevent excessive cognitive load and improve attention focus.
- HR-Recycler will develop a novel action planning to fulfill the robot's goal. *The autonomous robotic actions planning* will be enhanced with a dedicated controller to ensure robot safety, material safety, and human physical and perceived safety. HR-Recycler will advance the current state of the art in

[5] Confessore, G., Fabiano, M. and Liotta, G., 2013. A network flow based heuristic approach for optimising AGV movements. *Journal of Intelligent Manufacturing*, *24*(2), pp.405-419.

[6] Seif, R. and Oskoei, M.A., 2015. Mobile robot path planning by RRT* in dynamic environments. *International journal of intelligent systems and applications*, *7*(5), p.24.

robotic planning, by enhancing decision-making with a cognitive component (valence) in which action-selection is based on an ethics engine (T2.3).

- In HR-Recycler project a *trajectory planning method will be developed suitable for AGVs* that operate in a factory environment with human presence. This planner will be integrated with human factor aspects by developing human aware navigation capabilities. Human presence, motion intention and human comfort will be modeled and reported in the robot's map to regulate its navigation parameters in more strict safety range.
- HR-Recycler will build on the latest advances in AR-based HRC in order to develop a *multi-level system combining multimodal optical and projective AR.* The system will be developed according to cognitive ergonomics, improving attention and preventing excessive mental workload, so that it will enhance safety and increase productivity in the factory floor. In addition, HR-Recycler will perform extensive testing in real-world industrial settings with actual workers employed in the industry.
- Apart from a theoretical Trust Human Factor Worker model, in HR-Recycler project a trust *classifier will be developed*.
- This trust classifier will be based on a first experiment, oriented to detect which *psycho-physiological signals may be more correlated to trust/distrust situations* (section 4.1)
- To ensure that the appropriate action is selected by the robot, HR-Recycler will develop a novel decision-making model where the human co-worker's feedback will update and rebalance the robot's state-space when low confidence in the perceptual system is detected. The interaction (appropriate communication between the robot and the human) will be handled by T7.2 (Interaction Manager). HR-Recycler's *learning from human input* approach goes beyond the current state of the art in industrial collaborative robots, as it will improve decision-making and action-selection in cases of uncertainty where the robot could perform an incorrect action, or no action at all.

This deliverable is the first version and the final version, D7.2, will be released on Month 26 (January 2021). Considering all of this, the outcome is an initial specification of a trust factor model for human- robot collaboration.

1 Introduction

1.1 Overview

In Human-Robot Collaboration (HRC) we are dealing with in HR-Recycler is based on Human supervisory control of robots in performance of routine tasks. These include handling of parts on recycling disassembly lines and accessing and sorting electronical devices and components. Such robots can be both robotic arms or Automated Guided Vehicles (AGV), capable of carrying out a limited series of actions automatically, based on a computer program, and capable of sensing its environment and its own joint positions and communicating such information back to a human operator who updates its computer instructions as required.

Human-Robot Collaboration triggers emotions in humans. To classify and detect these emotions and responses can help to design a smoother collaboration and interaction among humans and robots.

When it comes to the detection and understanding of emotions, the role of the autonomic nervous system (ANS) is critical in the generation, expression, experience, or recognition of emotion. The responses of the autonomic nervous system, which is a general-purpose physiological system [1]**,** are actually quite specific, with different patterns of activation characterizing different situations and their associated emotional states. Emotional states are mediated by a family of peripheral, autonomic, endocrine, and skeleton-motor responses. The following table shows a relationship between visible ANS-mediated changes in emotions:

Type	Change	ANS-mediated Basis	Emotion
Coloration	Reddening	Vasodilation, increased contractility	Anger
	Blushing	Vasodilation	Embarrassment
	Blanching	Vasoconstriction	Fear
Moisture and secretions	Sweating, clamminess	Sweat glands	Fear
	Salivating, drooling	Salivary glands	Disgust
	Foaming	Salivary glands	Anger
	Tearing, crying	Lacrimal glands	Sadness
	Lubricating	Mucus membranes	Sexual arousal
Protrusions	Piloerection	Muscle fibers at base of hair follicles	Fear, anger
	Genital erection	Vasodilation	Sexual arousal
	Blood vessels bulging	Vasodilation	Anger
Appearance of eyes	Constriction	Pupils	Anger
	Dilation	Pupils	Fear
	Bulging eyes	Eyelid muscles	Anger, fear
	Drooping lids	Eyelid muscles	Sexual arousal
	Twinkling	Lacrimal glands plus contraction of orbicularis oculi	Happiness

Figure 1: Visible ANS-mediated changes in emotions – Levenson 2014 [2]

The inference of Human Factors is based on psycho-physiological measurements, integrating three different dimensions of emotional processes: experiential (e.g. feeling angry), behavioural (e.g. severe frown) and physiological (like EEG and heart rate). Here a summary of the different classical Human Factors detected on each of the dimension of emotion is presented, based on the main findings detailed in Section 3:

Dimension	***Human Factor Detected***	***Metrics***
Physiological	Attention (low/high), valence (positive/negative), arousal (exciting/calming), engagement and memory activation	EEG, HRV, GSR, respiration, EMG
Behavioural	Visual attention, valence (positive/negative) and arousal (exciting/calming)	Eye-tracking, facial analysis
Experiential	Valence (positive/negative)	Self-report

Table 1: Human factors and Emotions

Provided this brief introduction, the current challenges and research questions (what we want to achieve in HR-Recycler), we will answer in this document are the following:

- Which are the main ***Human Factors*** to have into account for a smooth Human-Robot Collaboration?
- What are the most appropriate ***psycho-physiological signals*** to use in Human-Robot Collaboration studies?
- How can we ***verify the applicability and accuracy*** of those psycho-physiological measures in Human-Robot Collaboration studies?

Furthermore, the KPIs related with this deliverable are the following:

- **KPI 1.1**: Tasks where robots interact verbally and non-verbally with humans
- **KPI 1.2**: Communication success rate (humans are able to correctly identify the messages of the robot and viceversa)
- **KPI 1.3**: Robot and human can maintain task-oriented dyadic interaction for completing goal state
- **KPI 5.2**: Percentage of human workers feeling safe working with the robots
- **KPI 5.4**: Increased speed of collaborative tasks, through improved human motion prediction
- **KPI 5.5**: AR toolkit fusing projective and optical see-through AR modalities to guide workers into HRC and further enhance worker perceived collaboration efficiency & safety
- **KPI 6.1**: Number of human factors to be considered for covering all the requirements from the pilots
- **KPI 6.5**: Delivery of a legal study on human-robot interaction and collaboration, acting as a guideline document of the ethical and legal requirements for conducting research about algorithm development through monitoring and assessing the actions of humans

1.2 Structure of the deliverable

The deliverable is structured as reported below:

Chapter 1 – Introduction provides a brief introduction to HRC, emotions related and psycho-physiological measurements to detect the different Human Factors (HF) related to Huma-Robot Collaboration.

Chapter 2 – Requirements for an effective communication identifies and considers the various factors that affect social interaction with robots, providing a detailed analysis of communication strategies between robots and humans.

Chapter 3 – Human Factors Worker Model presents an initial Human Factor Worker Model, beginning from a State-of-the-Art analysis of Human Factors related to Human-Robot Collaboration. The main findings of this SotA will be related to the three dimensions of emotional processes (experiential, behavioural and physiological). Additional to this fact, the specific HF related to trust in a successful Human-Robot Collaboration are also analysed. Together with this issue, a brief description of the psycho-physiological signals to be measured to detect those HF is done.

Chapter 4 – Pilot studies with end-users describes the two-level approach for the pilot studies. First of all, the ones for understanding the Trust Factor (Trust Game) and then the pilot studies for validation that will be conducted both in a Virtual Reality environment and in a controlled environment, inside the plant.

Chapter 5 – Assessment of robot social behaviours provides a set of pilot studies whose main purpose is to evaluate the interactive capabilities of robots. The goal of the proposed pilots is to assess which behaviours and means of communication allow robots to be accepted by humans as well as promote and improve collaboration between them.

2 Requirements for an effective communication between robots and human co-workers

Robots are created to perform a diversity of tasks, serve a variety of purposes and are extensively used in many domains. Traditionally, robots are employed in settings that require routine operations or are considered dangerous for humans. For example, operations like handling materials, assembling or painting are almost exclusively performed by robots [3]. These robots require some degree of autonomous operation as well as the ability to make decisions and perform tasks [4]. Until recently, most robot examples are constrained in situations where little interaction with humans is required. As nowadays the development of robots goes beyond utilitarian purposes, we observe a change in paradigm and robots that operate in close proximity to humans start to gain ground. And even if the primary goal of a machine is to perform a task that demands little interaction with humans (like maintaining a household clean), stable interactions still emerge [5]. It is, therefore, important to understand how these interactions emerge and what their nature is. This is one of the key goals of the field of Human-Robot Interaction (HRI), and what mainly defines the design and behavioural characteristics of a robot are mainly the application domain and the nature of the interaction. One of the most important challenges in the design of social robots (that is, robots that interact with humans) is to correctly identify and consider the various factors that affect social interaction [6][7].

In most HRI scenarios, it is widely accepted that autonomous and transparent behaviours are essential, as humans are capable of intuitively understand and interpret them. Thus, behaviours that resemble those of humans provide a more intuitive interface: if they fulfil the social expectations, they become predictable and interpretable [8]. Based on this approach, robots are bound to the social standards of Human-Human Communication and the social rules attached to the role they assume. Humans seem to intuitively apply the same social rules when they interact with machines as when they interact with other humans [9]. For example, theories from social psychology (and more specifically empathy and similarity) can indeed be transferred to HRI as more prosocial human reactions can be triggered [73][74]. For these reasons, a standard approach of HRI to study effective ways of communication that are applied in human-human interactions, and advocates of humanoid robots may claim that human-like design benefits HRI, as it enables communication channels that are similar to those of humans. Indeed, humans are highly communicative beings; they use a variety of nonverbal multimodal cues to communicate, such as gestures, gaze, facial expressions, prosody or even the way they move and their proxemics [10]. Humans are able to distinguish established trajectory planners from humans but not trajectories generated by human-like motion. Such studies suggest that human-like motion behaviour enhances the acceptance and collaboration between robots and humans [75]. These rich nonverbal communication channels contribute to the enhancement of meaningful cues that can complement or even substitute spoken dialogue, like in the context of behaviour authoritativeness [11].

When it comes to collaboration, effective communication is essential for the successful completion of a task. By collaboration we mean “the mutually beneficial and well-defined relationship of two or more entities to achieve a common goal” [11]. In these terms, a team is formed when two or more entities (or agents) that have complementary skills perform common tasks and share common goals. To achieve effective collaboration, effective communication within a team is essential for the coordination of teamwork, especially since there is a potential divergence of mental states between the team members [13]. For example, dialogue is key for the formulation and execution of shared plans [14].

In Human-Robot Collaboration (HRC) settings, the team is mixed and typically comprises humans and robots working together. For collaboration to be efficient, robots are required to robustly perform a task, be trustworthy and effectively communicate with the human co-worker. Additionally, we highlight the importance of safe operation, as the robot will be required to function in proximity to humans. In the context

of HR-Recycler, we will be employing non-humanoid robots that are classified as robotic arms or AGVs. Here, these robots will be responsible not only to successfully perform the tasks that are assigned to them but also effectively collaborate and communicate with their human co-workers. Typically, most communication approaches in the field of HRI employ the anthropomorphist approach: they use humanoid robots that act as similar to a human communication interface. In the HR-Recycler project, our approach is different, as the design of the robot serves the task of device disassembly; the proposed robots are suitable for efficient grasping and device manipulation and not for interaction, as they do not have access to an anthropomorphic communication channel.

In the field of HRI we have not found extensive studies to deliver novel ways of interaction suitable for non-anthropomorphic robots. The HR-Recycler project seeks to address the gap in communication for non-anthropomorphic robots, based on theories drawn from human communication [15]. Both Human-Human Interaction (HHI) and Human-Robot Collaboration (HRC) domains require knowledge about the principles and characteristics of social interaction. On the one hand, HHI can benefit from HRI experiments as studies with social robots can act as the testing ground for theoretical explanations. On the other hand, HRC can benefit from studies of HHI, as information acquired from this domain will inform adaptive systems that allow robots to fluently interact with humans. HR-Recycler will advance the field of HRI and HRC by including mechanisms that underlie social competence in a broader range of non-human social behaviours for communication and collaboration.

2.1 Analysis of communication strategies between robots and humans

In order to successfully analyse communication strategies between robots and humans, we first need to understand communication in social interactions between humans. As mentioned previously, humans employ not only verbal but also non-verbal multimodal cues to communicate that can complement or even substitute spoken dialogue. Non-verbal communication can be *explicit* (where communication is deliberate, and the sender has the goal of sharing information) such as pointing at an object, and *implicit* (information is not deliberately communicated but is conveyed by inherent behaviour) such as gaze direction. Both forms are important for successful communication. For communication to be effective, the message needs to be clear, and the channel of communication interpretable by the receiver. Additionally, context is of high importance for a message to be clear, and several parameters can affect communication. Although in most collaborative tasks, verbal communication seems the most appropriate solution, in most industrial settings, the environment is noisy, so verbal communication is not as effective. Here, non-verbal communication seems to be more suitable, and the challenge is to correctly convey information with non-humanoid robots.

In the following sections, we provide further details regarding existing communication strategies employed in HRI and HRC, what are the key variables that affect communication and what is the project's approach.

2.1.1 Communication objectives

In the context of collaboration, the key objective of communication is to share information that is important for the successful completion of a task. According to Levinson [15], successful interactions are bootstrapped in the intrinsic motivation to interact and communicate. More specifically, communicative capabilities include the motivation to interact and communicate through universal (language independent) manners, like looking at objects or pointing at them. For Levinson, humans are natively endowed with a set of cognitive abilities and behavioural dispositions that work together to endow human face-to-face interaction with certain special qualities. In addition, humans seem to respond to actions or intentions and not behaviours. In communication, actions are generated taking into account that they will be interpreted by a specific other;

thus, actions have to be designed to be transparent, and most actions can be inferred regardless of language. Communication is by large cooperative, as those who wish to communicate a message intent that their actions are interpretable, however, it is characterized by expectation of closed timing: an action produced in an interactive context sets up an expectation for a timely response. Intention attribution is important as humans infer likely goals that would have motivated the behaviour: the communicated message is interpreted and the reasoning behind the message inferred.

In HR-Recycler, we have identified the following situations where communication is needed:

- Information about the task at hand that includes the general goal, which are the human co-worker's and the robot's tasks and the steps needed.
- Error detection such as action failures, low confidence in object or human identification, and missing steps or wrong categorization of a component during the disassembly.
- Risk detection which mainly includes raising awareness of the situation and alerting the human co-worker (human, robot or object safety hazards).

2.1.2 Worker and Human Factors affecting communication

In industrial settings, safety and risk detection and assessment is of high importance for both the physical safety of the human worker as well as the successful completion of a task. Risk detection does not only include the physical but also the mental safety (or perceived safety) of the user. Different users have different approaches when it comes to closed interactions with robots and stress or anxiety can be induced by these interactions. By measuring implicit information to evaluate the internal state of the user (such as relevant physiological parameters like heart rate or galvanic skin response), we can enhance the interaction by monitoring and interpreting this implicit feedback. For example, if we detect certain changes in the physiological reactions of a user that imply stress when a robot comes in close proximity to them, we can improve the interaction by ensuring that the robot's proxemics allow the user to feel comfortable.

Further factors that may affect interaction and communication is cognitive load. There are cases when a task may be overwhelming for the mental capabilities of a user. High cognitive load may not only affect the successful completion of the task but also effective communication between peers, as information will not be processed appropriately. Here, the behaviour of the robot can alleviate the cognitive load of the user by adapting its behaviour or the task to the needs of the worker. This approach, that is, adapting the behaviour of the robot according to the internal state of the user is called affective robotics and it has been extensively applied in social robots (such as humanoids), however, this approach is not common in industrial practice. In HR-Recycler we introduce affective robotics in industrial settings and enhance HRI by taking into account human's internal states and cognitive information processing.

2.1.3 Channels for multimodal communication

In terms of explicit communication, speech is the most common channel employed and, in several cases (such as procedural instructions), the message is clear and correctly received. Communicative gestures are used to inform peers about intention or call attention to an object or location. Overall, most gestures seem to be easily understood, especially if they are primitive signs, such as stop or turn. Facial expressions are used to communicate the internal state of a peer. Haptic feedback is another form of explicit communication as two or more entities are connected either directly (like a handshake) or through an object they both manipulate.

Implicit communication can be more challenging than explicit communication, though equally informing at times. In implicit communication, the partner needs to derive the intention (or internal state) of others by observing them. In this case, context and action recognition are pivotal. Examples of implicit communication include gaze direction, as information about the object or location of interest can be extracted from gaze. Other forms of implicit communication are the usage of means to extract physiological signals from human partners that provide information about the internal state of the other. Here, brain or muscle activity signals changes in thoughts or actions whereas heart rate or galvanic skin response can inform the stress levels or emotional state of the other, even if such information is not communicated explicitly.

In industrial settings, many of the explicit or implicit communication channels cannot be easily employed. For example, speech may not be possible due to loud noises coming from the operation of machinery. Workers usually tend to wear protective masks and head gear (like glasses or helmets), so facial expression recognition or even sometimes gaze direction are not possible. Gestures, on the other hand, can be recognized more easily compared to gaze or facial expressions. Such difficulties, along with the use of non-humanoid robots, raise certain challenges in finding solutions that foster and facilitate interaction without having workers specially trained to interact with robots. As workers will be performing tasks with robots in a collaborative way, the introduction of intuitive interfaces for communication is key to ensure that workers concentrate on the task at hand and not on how to appropriately communicate. A widely used interface for HRI in industrial settings is the application of Augmented Reality (AR) and Virtual Reality (VR). Such approaches may increase productivity and at the same time, enhance human safety [16]. AR may be preferred to VR if workers need to maintain a sense of the real world. In contrast, VR is preferred if workers need to be immersed in the virtual environment. AR technologies are usually preferred, as they allow humans to act in physical space and interact with physical objects while at the same time, they provide digital information or signals to draw attention that humans would not have access otherwise. They are inherently safer than VR as humans are aware of their environment, and if the information is presented appropriately, they are not distracting humans from performing the task but actually enhance productivity.

Examples of Augmented Reality applications in industrial settings have been used in assembly planning [17], guidance [18] or robot trajectory planning [19]. As mentioned in [16], most of these approaches have been tested in controlled laboratory environments. HR-Recycler further contributes to HRC by bringing these channels of communication to the actual industrial settings where the communication frameworks will be used and evaluated by real workers.

2.1.4 Humanoid vs common robot comparison in communication

In general, social interaction does not necessarily require a body to be successful. In physical collaborative tasks, the embodiment is important not only for task execution but also for perception and action. Embodiment is the inherent link to intelligence and provides qualitative advantages over non-embodied interfaces, given that the body is used to leverage the knowledge of human communicative behaviour [20], which in turn improves information transfer [21]. Embodiment, combined with shared context and physical presence, is critical for establishing successful communication. For example, embodied robots are perceived as more trustworthy compared to non-embodied ones [22] and were evaluated as effective communicators [23].

Regarding the robot's morphology and design, a rule of thumb is that its design should serve the task it was meant to perform. When it comes to interaction, the design of the robot may affect believability and acceptability. Studies suggest that the physical appearance of a robot biases the interaction as it may affect users' perception [24] and expectations about its social capabilities. For example, if a robot has eyes, it is expected to see; if it has hands, it is expected to manipulate objects. To be effective, a robot does not always

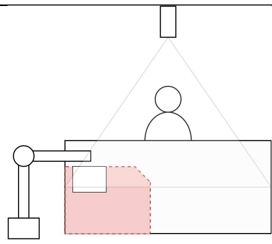
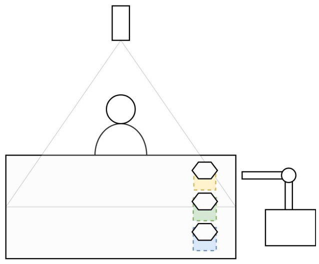
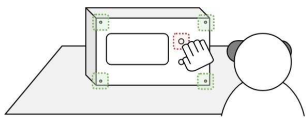
need to resemble a human. Advocates of humanoid robots claim that human-like design benefits and facilitates HRI, as such morphology enables communication channels that resemble those of humans. Indeed, an essential requirement for successful communication is the readability and transparency of the employed channel. In general, the disposition is for humanoid robots to be preferred by users [25]. Preferences in interaction and communication, however, are also task-related. Machine-like or human-like appearance may influence the perceived robot personality [25] and responsibility assumed by humans in a collaborative task: humans feel more responsible when interacting with machine-like robots compared to more anthropomorphic ones [26].

Although the industrial robots employed in the context of HR-Recycler do not fulfil the humanoid approach and take advantage of the communication channels that are employed by human-like bodies, they do satisfy the task they were meant to conduct. Natural, transparent and effective communication may be more challenging compared to that with humanoids; nonetheless, in HR-Recycler, we will explore novel and intuitive ways of delivering natural and transparent communication and interaction.

2.2 Dynamics in the interaction between humans and robots in HR-Recycler

Human-Robot interaction in the context of the HR-Recycler project is defined in accordance to the task performed, as well as in accordance to the interaction medium at use (physical, AR). In order to satisfy safety requirements, interaction through AR is generally advocated, however there exist also instances where physical collaboration exists, such as handing off components by the robotic manipulators to the human workers. Table 2 outlines and categorizes the identified collaboration scenarios.

	Occurrence	Relevance	Subject	Description
Physical Human-Robot Collaboration				
1	Regular	Process	Hand off components to be disassembled by the human worker	The fixed robotic manipulator hands off the components to be disassembled by the human worker to appropriately appointed locations on the worker side of the workstation.
2	Regular	Process	Retrieve and replace material containers on the workstation	The mobile manipulator on the AGV will retrieve and replace the containers containing classified materials and components from/to the disassembly workstation and transport them to the sorting machines.
AR-Enabled Workstation Collaboration				
1	Regular	Safety	Mark workstation plane with inbound robot manipulator trajectories by projection	Users perform disassembly on a workstation adjacent to the robot workstation, however the compliant manipulator will often reach to transfer device parts for further disassembly by the user. Projective AR will mark out the areas to be occupied by the manipulator and indicate the area that is safe for the worker to continue working. This will be complemented with human-aware manipulator trajectory planning to potentially alter robot trajectories in case of violation or a stop action mechanism to ensure human safety.

				
2	Regular	Process	Annotate item locations on the workstation plane for AGV pickup	Through the use of projective AR or AR glasses the locations of appropriate bins for placement of extracted device components by the user will be annotated, and the user will be warned in case of mistakes in placement (e.g. placing component in the wrong bin) 
3	Exceptional	Process	Indicate features with low confidence and allow user to confirm	In cases where device features relevant to a specific action (e.g. screw slots while unscrewing FPD cover) cannot be detected with confidence by the HR-Recycler system, the process will pause, and the worker will be notified to assist. Through the use of AR glasses, relevant detected features will be indicated, and the user will be asked to confirm them. 
4	Exceptional	Process	Notify user to locate and undo undetectable fixture	In case of device fixtures that are not possible to be detected or manipulated by the robot, the process will pause, and the user will be notified to assist. Through the use of AR glasses, the system will indicate the relevant device parts (e.g. device cover) and, if detected, the potential location of the fixture. The user will be asked to identify and remove the fixture.

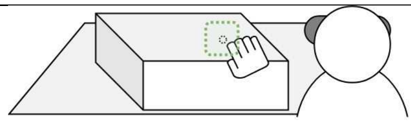
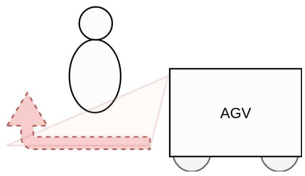
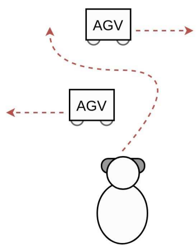
				
AR-Enabled Factory Floor Collaboration				
1	Regular	Safety	Indicate AGV path through projective and auditory cues	During movement, AGVs will indicate the intended trajectory through projection of immediate path on the ground plane. In addition, auditory signals will warn users whose projected movement intersects with that of the AGV. This will be complemented with a safety mechanism that stops the robot's navigation in the case that a human intersects the trajectory or does not respond to the warning signals. 
2	Regular	Process	Indicate optimal path to user	Indicate, through visual cues presented through the use of AR glasses, the optimal path to take in order to reach a target location, e.g. move from one workstation to another, minimizing intersections with AGV paths and traveling distance. 

Table 2: HR-Recycler use cases and interactions between human and robots

This summary will serve for defining the best pilot studies related to the different use cases for being tested.

3 Human Factors Worker model

3.1 State-of-the-art analysis of Human Factors related to Human-Robot Collaboration

The state-of-the-art analysis presented herein targets the improvement of our knowledge in understanding and modelling human factors for Human-Robot Collaboration. This knowledge is expected to significantly facilitate the development of a reliable, robust and impactful Human Factor Worker Model while answering the current challenges and research questions defined in Section 1.

3.1.1 Literature Search

We searched the bibliographic databases of Science Direct, Springer Link and Scopus to identify Human Factors which have to be taken into account in Human-Robot Collaboration, published after 1999.

The aim was to identify different human factors and emotions that can affect to a positive Human-Robot Collaboration. This HRC is mainly applied to the human supervisory control of robots in performance of routine tasks. These include handling of parts on manufacturing disassembling lines, but also accessing and classification of different types of WEEE. Although there are differences between the concepts of Human-Robot Collaboration (HRC) and Human-Robot Interaction (HRI) [27], interaction is a more general term with includes collaboration. HR-Recycler focuses on how human and robot can work together for reaching a common goal so we will use both terms as equivalent.

The inclusion criteria for study selection were the following: a) the study must be conducted with adults, either in a real-life research setting or retrospectively, b) experiments for detecting human factors and the method used should be described, c) quantitative outcomes of the study should be presented, and d) the paper describing the study, must have been written in English. These criteria allowed us to find research papers focused on emotional process and non-technological perspective. In order to focus the search on the project domain, we performed another search adding an additional criterion: e) the study should focus on HRC. Ongoing studies, case reports, surveys or reviews, qualitative studies, studies describing protocols, and all studies conducted before 1999 were excluded from the review. Besides, we distinguish between research papers from technological and non-technological.

The first search within the title, abstract, and keywords of the manuscripts was performed using the following terms: *"Human Factors" OR "emotion models" OR "emotional process"*. For the second search focused on HRC, a search within the title, abstract, and keywords of the manuscripts was performed using the following terms: *"Human Factors" OR "physiological signal" AND "Human-Robot Collaboration".* The whole process was based on the PRISMA guidelines [28].

3.1.2 Quality Assessment of Papers

After proceeding with the two searches detailed in previous section, 49 articles were downloaded, and when read and analysed, a total of 11 relevant articles were found. Each article was then deeply reviewed and classified according to the search criteria: more generic for Human Factors and emotion model (HF & EM) and more specific for Human Factors and Human-Robot Collaboration (HF & HRC) (see Annex 1 for details).

3.1.3 Main findings

Our state-of-the-art analysis examined the current status of human factors detection and monitoring at two level: first at general level and, then applied in HRC domain. The main outcome of this analysis is that the Human-Robot Collaboration triggers some emotions in the human, which impact on how the human interacts with the robot. To quantify and detect these emotions can help to design a smoother collaboration and interaction among humans and robots.

In next subsection we provided a detailed description of this analysis. First, we provide an analysis of the different dimensions of an emotional process and which are the main psycho-physiological signals to measure these processes. Then we explained the classical Human Factors which can be inferred by using those psycho-physiological signals and also how trust is a key element is HRC.

3.1.3.1 Dimensions of emotional processes

The inference of Human Factors is based on psycho-physiological measurements, integrating three different dimensions of emotional processes:

- experiential (e.g. feeling angry),
- behavioural (e.g. severe frown) and
- physiological (like EEG and heart rate).

The following picture shows a summary of each dimension, related signals and human factors; and also, the advantages and drawbacks of each measurement method.

Psychophysiological dimension	***Signals***	***Human factors/Metrics***	***Advantages***	***Drawbacks***
Physiology	Electrical activity of the brain: Electroencephalography (EEG), magnetoencephalography	Valence *(positive/negative)* Arousal *(exciting/calming)* Engagement Memory activation Cognitive load	– Objectivity: assessments don't depend on user's perception – They can be used in multidimensional measures to provide different faces of user's state	– Inconsistent physiological reading among participants → define a baseline – Devices can cause more stress to participant – Sensors are sensitive to environmental noise
	Cardiac function: Heart Rate Variability (HRV), electrocardiography (ECG), plethysmograph	Arousal *(exciting/calming)* Cognitive load	– Cheap – Easy to setup and use	– Sensors are sensitive to environmental noise
	Conductivity of skin: Galvanic Skin Response (GSR)	Arousal *(exciting/calming*	– Easy setup and to be understood – Cheap	– Response lags stimulus – Sensitive to motion – Requires baseline
	Blood Volume Pulse (BVP)	Valence *(positive/negative)*	– Quickly collected and analyzed – Very simple device to collect data – Combined with eye-tracking can be a powerful tool	– Sensors are sensitive to environmental noise
	Respiration	Valence *(positive/negative)* Memory activation Cognitive load		– Under both voluntary and involuntary control – Obtrusive to collect – Not as worth as other physiological signals
	Muscle electrical activity: electromyography (EMG), facial EMG	Valence *(positive/negative)*	– Sensitive to subtle signals	– Facial hair can interfere
Behaviour	Eye-tracking	Visual attention	– Provide additional findings to self-report – Good metric for interface evaluation and usability testing	– Data are easy to manipulate. How? Changing the range of fixations (there is no standard) – Focus vs locus of attention ("do you really pay attention at what you see?"
	Pupil Dilatation	Arousal *(exciting/calming)* Cognitive load	– Non-contact required	– Requires controlling the brightness of the environment
Experiential	Geneva Emotion Wheel PrEmo	Valence *(positive/negative)*	– Very intuitive and easy to use	
	LEMtool	Valence *(positive/negative)*	– Offers 8 "web-relevant" emotions – Cross-cultural emotional self-report tools	– Participants not only rate the web pages on visual appeal, but can also be influenced by texts and images

Figure 2: Summary of psycho-physiological signals and related human factors

3.1.3.2 Main psycho-physiological signals to measure emotional processes

A brief description of the psycho-physiological signals used for measuring emotional processes mentioned in previous section is provided here.

- **Heart Rate Variability (HRV)** is a signal to indicate perceived safety / danger in the environment. This physiological mark is measured by the length of consecutive inter-beat intervals (RR-intervals). The features assigned to HRV are related to frequency: low LF-HRV (from 0.04 to 0.15 Hz) and high HF-HRV (from 0.15 to 0.4 Hz). Primary methods for tracking heartbeats are ECG and blood pressure and the pulse wave signal extracted from a photoplethysmography (PPG).
- **EEG signal (Electroencephalography)** captures cortical activity of the brain, exhibiting changes in human thoughts, actions and emotion. it is appropriate to ***use in the emotional state recognition tasks***. It is the brain activity in response to specific event: Event-related potential (ERP). The features extraction is related to the frequency bands associated to brain activity, denominated as rhythms, and several studies suggest that are associated to specific mental tasks [34].
- **GSR (Galvanic Skin Response)** is the conductivity of the surface of the skin. It provides a measure of the resistance of the skin (electrodermal activity) by positioning two electrodes on the top of two fingers. This resistance decreases due to an increase of sudation, which usually occurs when one is experimenting emotions such as stress or surprise. Since sweating is out of control of the human, it is regulated by the autonomous nervous system (ANS). ***High arousal*** of the sympathetic branch of ANS increases the sweat gland activity, which accordingly increases skin conductance and another way around. This is the reason why GSR can be considered as ***an index of emotional and sympathetic responses***. The feature extraction for GSR is done through the decomposition of the phasic (fast-changing) and tonic (low-changing) components of the skin's responses for: Maximum deflection, in net signal, maximum phasic component, net phasic component. In this sense, mean value of the GSR is related to the level of **arousal [35]** and stress is associated with trust and cognitive level. High stress levels are associated with high cognitive levels and low trust.

3.1.3.3 Classical Human Factors inferred by psycho-physiological signals

As described in Figure 2, there are five classical Human Factors that the psycho-physiological signals provide.

- **Valence** is a term from psychology that refers to the level of pleasantness/unpleasantness when interacting or observing an object, event or situation. Valence could be described by polar scales that, in aggregate, defined a continuous dimension from pleasantness (happy, pleased, hopeful, etc) to unpleasantness (unhappy, annoyed, despairing, etc) [29]. It is related to the positive/negative information that the person is processing related to the object, event or situations that people experience. Valence dimension is present in virtually all systems that have been developed to classify emotion.
- **Arousal** is a state of heightened physiological activity. This includes having strong **emotions** like anger and fear and we go to the **emotional arousal** state in response to our daily experiences. For example, the fight, flight or freeze response is a state of **emotional arousal.**

  These two Human Factors (Valence and Arousal) are called "core affect" by Russel [30], that's to say a neurophysiological state that is consciously accessible as a simple, nonreflective feeling that is an integral blend of hedonic (pleasure–displeasure) and arousal (sleepy–activated) values specifically identified dimensions of pleasant–unpleasant, tension–relaxation, and excitement–calm as the basis of feeling and emotion.

In factor analysis and multidimensional scaling studies, emotional valence is one of two axes (or dimensions) on which an emotion can be located, the other axis being arousal (expressed as a continuum from high to low). For example, happiness is typically characterized by pleasant valence and relatively high arousal, whereas sadness or depression is typically characterized by unpleasant valence and relatively low arousal.

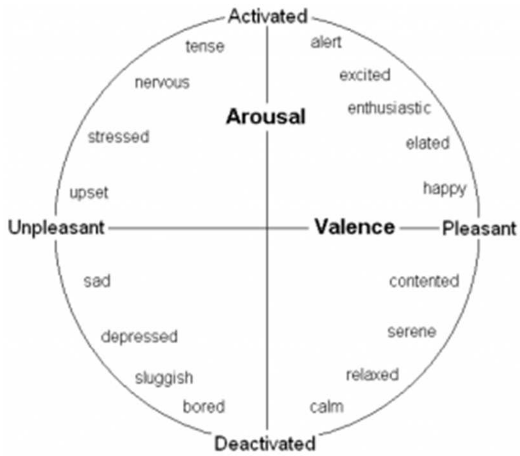


Figure 3: Circumplex model of emotion based on Valence and Arousal [31][32]

- **Cognitive load,** which refers to the total amount of mental activity imposed on working memory in any one instant. The memory function is in the following way:
  - o We build schemas in working memory and integrate them into existing schemas in long-term memory
  - o We bring schemas from long-term memory into working memory to understand the world

  But, when the working memory must process too much information, it leads to poor comprehension and it obstructs learning.

  One cause of too much demand on working memory comes from an abundance of novel information (more information than the person can process). But high cognitive load is also strongly influenced by the number of elements in working memory that interact with each other (the greater number of elements, the more cognitive workload).

- **Memory activation** refers to the degree of intensity of cognitive processes. It captures the degree of storage, encoding and retention in memory. After this memory activation, **r**ecall or retrieval of memory refers to the subsequent re-accessing of events or information from the past, which has been previously encoded and stored in the brain. During recall, the brain "replays" a pattern of neural activity that was originally generated in response to an event, echoing the brain's perception of the real event. There are **two main methods** of accessing memory:
  - o **Recognition** is the association of an event or physical object with one previously experienced or encountered and involves a process of comparison of information with memory, e.g. recognizing a known face, true/false or multiple-choice questions, etc.
  - o **The recall** involves remembering a fact, event or object that is not currently physically present (in the sense of retrieving a representation, mental image or concept), and requires the direct uncovering of information from memory, e.g. remembering the name of a recognized person, fill-in-the blank questions, etc.
- **Engagement:** Psychologists have defined engagement as a sort of ongoing emotional, cognitive and behavioural activation state in individuals. In this sense, engagement also refers to involvement,

commitment, passion, enthusiasm, absorption, focused effort, zeal, dedication, and energy. In a similar vein, the Merriam-Webster dictionary describes the state of being engaged as “emotional involvement or commitment” and as “being in gear” [33].

All these five Human Factors have been classically used to measure and understand emotional processes. However, some higher-level cognitive processes such as trust play a key role to achieve a successful Human-Robot collaboration. Several studies like [84] and [85] show that the five classical human factors are related and can be used as trust metrics. Useful as this approach may be, doing so implies transforming the direct biometrical signal into an intermediate indicator in order to predict the desired feature, thus creating significant delays, decreasing prediction accuracy and making it more sensible to external measurement perturbations, which may be quite significant on an industrial environment. Therefore, it is necessary to explore other alternatives that could lead to employ biometrical signals as direct indicators of trust.

Next sections explore the concept of trust in order to understand its nature and dimensions and provide a list of different factors affecting the HRI. Section 3.1.3.7 provides a summary of the current work regarding the use of psycho-physiological signals to identify human trust, where has been identified a relevant lack of formal studies in this area. All these concepts will establish the basics regarding how human trust works and how could be the best way to measure its variations; thus, they will play a vital role in order to define the experimental process to be committed on the project.

3.1.3.4 One step further: multidisciplinary nature of trust

Trust is a major factor to determine effective collaboration among agents, both human-to-human and human-to-machine, alike. Thus, researches involving trust modelling and measuring involve several disciplines as psychology, sociology, biology, neuroscience, economics, management, and computer science [36]-[45]. Even if both approaches – modelling and measuring – are related and thus share common knowledge, each has a unique purpose and identity and, thus, they require different considerations. While trust modelling aims to represent the nature of human trust behaviour extrapolating individual responses to a universal level, trust measuring looks for mensurable involuntary body responses when the subjects are exposed to different trust-related stimuli.

As a general basis, trust can be understood as a relationship in which a subject (trustor) interacts with an actor (trustee) under uncertain circumstances in order to achieve an expected goal.

However, due to the broad variety of disciplines involving the study of human trust, it is complicated to define a unique modelling approach. The reference survey [36] provides a deep insight of several historical analysis of several multidisciplinary approaches towards the concept of trust and how it is described by the different fields of interest. These variety of definitions are shown on Table 3.

Discipline	Meaning of Trust	Source
Sociology	Subjective probability that another party will perform an action that will not hurt my interest under uncertainty and ignorance	[37]
Philosophy	Risky action deriving from personal, moral relationships between two entities	[38]
Economics	Expectation upon a risky action under uncertainty and ignorance based on the calculated incentives for the action	[39]

Psychology	Cognitive learning process obtained from social experiences based on the consequences of trusting behaviours	[40]
Organizational Management	Willingness to take risk and being vulnerable to the relationship based on ability, integrity, and benevolence	[41]
International Relations	Belief that the other party is trustworthy with the willingness to reciprocate cooperation	[42]
Automation	Attitude that one agent will achieve another agent's goal in a situation where imperfect knowledge is given with uncertainty and vulnerability	[43]
Computing & Networking	Estimated subjective probability that an entity exhibits reliable behaviour for particular operation(s) under a situation with potential risks	[44]

Table 3: Multidisciplinary definitions of Trust [36]

These definitions offer some topic-specific insights than enhance the early description of trust in order to consolidate a more detailed definition of the concept:

"Trust is the willingness of the trustor (evaluator) to take risk based on a subjective belief that a trustee (evaluatee) will exhibit reliable behaviour to maximize the trustor's interest under uncertainty (e.g., ambiguity due to conflicting evidence and/or ignorance caused by complete lack of evidence) of a given situation based on the cognitive assessment of past experience with the trustee." [36]

This definition points out several critical issues regarding the nature of trust:

- *A subjective belief:* The perception of trust depends highly on the interacting individuals and the preconceived image of who other behave has a great impact on it.
- *To maximize the trustor's interest:* In order to detect to influence of trust between trustor and trustee, the interaction must imply some profit or loss for them that vary according to the trust/distrust interactions.
- *A reliable behaviour […] under uncertainty:* The trustor bases his interaction according to an expected behaviour from the trustee; thus, he interacts in order to maximize the expected outcome according to his own interests, even if this implies a suboptimal outcome. This means that the capacity to foresee the outcome of the interaction and act in consequence yields to trustier scenarios. In other words, trust is more related to the reliability and uncertainty of the response rather than to its nature.
- *Cognitive assessment of past experience:* Even if the prior stages of a trustor-trustee pair may be strongly influenced by the preconceived subjective beliefs, future interactions provide a positive/negative insight of how each actor behaves when they interact among each other. This means that trust possesses a dynamic nature.

3.1.3.5 Dimensions of trust

Several authors refer to different dimensions of trust to describe different aspects of the origin of trust. For instance, [45] distinguishes a *moralistic trust* based on the previous beliefs about others' behaviour and *a strategic trust* based on the individual experiences. Hoff and Bashir [46] identify three different categories of trust: *dispositional* (based on individual features such as culture, gender and age), *situational* (circumstantial factors such as task difficulty or criticality that may modify the otherwise nature response) and *learned trust* (based on the accumulated experiences). Every interaction between trustor and trustee is a combination of these three facets. Similarly, [47] describes trust as a combination of *phenomenon-based trust* (based on the trustor's subjectivity and his perceived social attitude), *sentiment-based trust* (related with the need to trust or not the trustee and the consequence if the decision fails) and a *judgement-based trust* (liked to how the trust is updated with new interactions).

On the other hand, [36] considers trust as a combination of *individual trust* (derived from own personal characteristics and conformed by *logical trust* and *emotional trust*) and *relational trust*, referenced to the dimensions of trust that rise from the relationship with other entities. Even if different authors use different classifications, it is possible to roughly map different similarities between different approaches. Even if some minor particularities of different dimensions are omitted, Table 4 shows how they roughly come together. For convenience, we will use the nomenclature defined in [46] – dispositional, situational and learned trust – without loss of generality.

<table>
<tr><th colspan="4">Similarity map of trust dimensions according to different authors</th></tr>
<tr><td>[45]</td><td>[46]</td><td>[47]</td><td>[36]</td></tr>
<tr><td rowspan="2">Moralistic</td><td>Dispositional</td><td>Phenomenon-based</td><td>Emotional</td></tr>
<tr><td rowspan="2">Situational</td><td rowspan="2">Sentiment-based</td><td rowspan="2">Relational</td></tr>
<tr><td rowspan="2">Strategic</td></tr>
<tr><td>Learned</td><td>Judgement-based</td><td>Logical</td></tr>
</table>

Table 4: A rough similarity map between trust dimensions among authors

3.1.3.6 Controllable trust factors within Human Robot interaction

It is important to note that different dimensions may or may not be controllable. Dispositional trust presents almost no controllability. On the other hand, even if the personal responses vary from one individual to another, it is possible to act over the factors conditioning the situational trust or even the evolution of learned trust in order to enhance the level of trust. Several authors have studied the influence of different factors in Human-Robot interaction from different perspectives, but [48] provides a deep review of the most relevant

factors across the bibliography, classified according to the Operator-Robot-environment triad. Figure 4 identifies these factors and links each of them to the most relevant trust dimension.

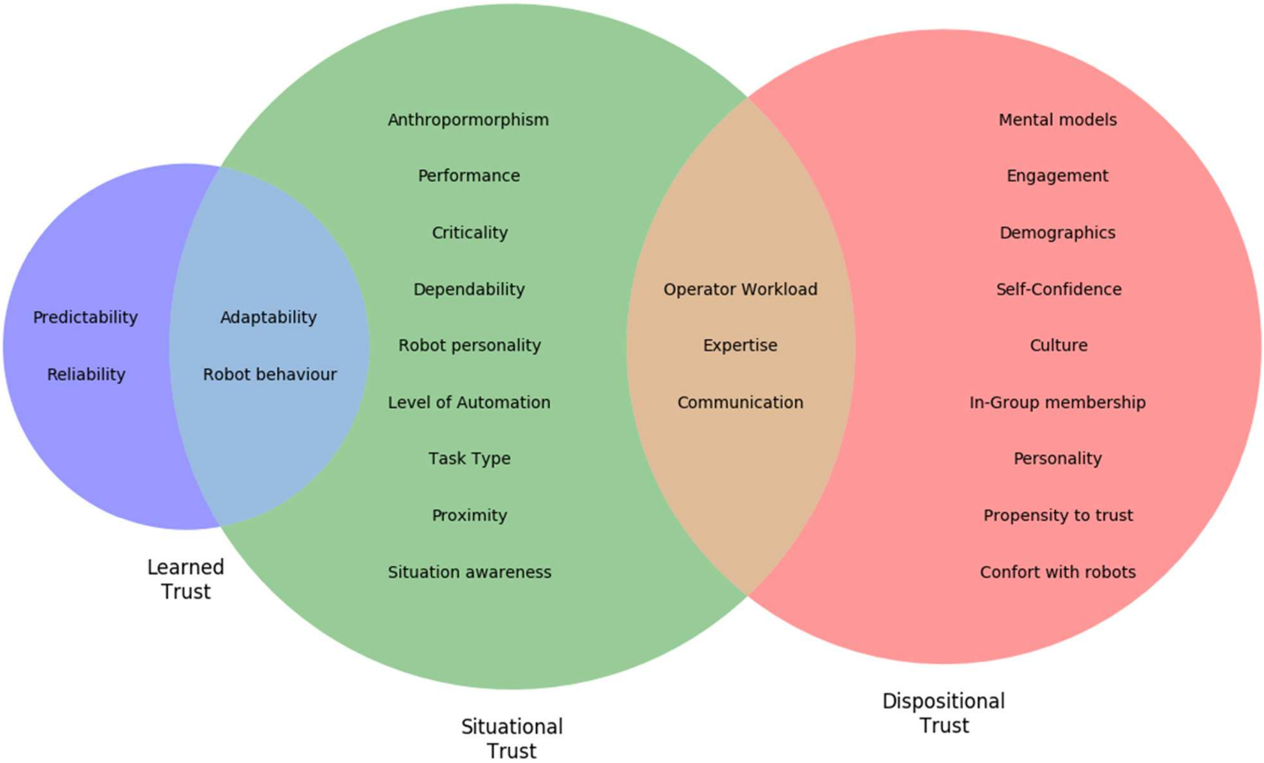


Figure 4: Classification of HRI trust factors

Figure 4 identifies the 26 trust factors mentioned in [48]. However, it is necessary to detect those that may result more relevant to the needs of the HR-Recycler project. [49] offers a review of these same factors, describing which of them have demonstrated experimental or correlational analysis. Furthermore, pure dispositional factors are intrinsic of the personal nature and, thus, are not controllable by any means. Only factors within the situational or learned trust dimensions will be considered with a demonstrated correlation with trust with be considered as an early approach. These factors are further detailed in the following paragraphs, along with a brief description of related experiments:

- **Robot behaviour**: Dautenhahn discusses the HRI from a social perspective, analysing the necessary social skills for robot companions[7], claiming that they should be considerate, proactive, non-intrusive and flexible. A truly personalized robot companion takes into consideration an individual human's likes, dislikes and preferences and adapts its behaviour accordingly [50][51]. In the manufacturing field, authors of [52] vary the robot performance and speed to accommodate to the muscular fatigue of the human operator and analyse the trust variation using time-series.
- **Reliability**: In the experiment performed in [53], participants were presented with a series of testing trials in which they diagnosed the validity of a system failure using only information provided to them by an automated diagnostic aid with varying reliability. Both subjective measures and objective

[7] "A robot companion is a robot that (i) makes itself 'useful', i.e. is able to carry out a variety of tasks in order to assist humans, e.g. in a domestic home environment, and (ii) behaves socially, i.e. possesses social skills in order to be able to interact with people in a socially acceptable manner." [50]

measures of performance were examined. Results show that objective performance measures of reliability were related to, but not perfectly calibrated with, subjective measures of confidence and reliability estimates (trust).

- **Level of automation**: In [54] participants are committed to perform certain complexity and difficulty variable actions with the aid of an automated system. These included both Sheridan-Verplank level 6 and 7 systems. Results show that variations in the automation levels according to task complexity and difficulty cause variations in operator's trust.
- **Proximity**: [55] explores how a robot's physical or virtual presence affects unconscious human perception of the robot as a social partner. Subjects collaborated on simple tasks with either a physically present humanoid robot or a video-displayed virtual robot. Study shows that many subjects interacting virtually had greater difficulty to accurately complete the designed tasks than those interacting with a physical robot.
- **Adaptability**: [56] compares team performance outcomes for robots controlled by *Chaski*[8] with robots that are verbally commanded, step-by-step by the human teammate. Trust was evaluated by a Likert questionnaire, showing that people in the first group agreed with statement "the robot is trustworthy," more strongly than people in the second group.
- **Anthropomorphism**: A full study regarding robot anthropomorphism and its influence in human trust is related in [57]. Results showed that anthropomorphic agents show greater trust resilience (a higher resistance to breakdowns in trust) and that incorporating human-like trust repair behaviour largely erased differences between human and robot agents. Automation anthropomorphism is therefore a critical variable that should be carefully incorporated into any general theory of human–agent trust.
- **Communication**: [58] studies who transparency in robot to human communication may influence in the level of trust. Participants were told to interact with different transparency-level robots to perform reconnaissance missions. Results show that higher transparency is more likely to increase human trust when the existing trust level is low, but it is more likely to decrease it when it is already high. Similarly, [59] states that a more detailed communication during medical human-robot interaction has a consistent effect on people's trust and perceived comfort.
- **Task type**: In [60], researchers investigate the effects of culture, robot appearance and task on human-robot interactions. Result show strong and positive correlations between interaction performance (active response and engagement) and preference (likeability, trust and satisfaction) in the human-robot interaction. However, if only the influence of task variability is considered, the study indicates that participants had higher active response in the tasks with higher sociability than in the task with low sociability, but show no significant correlation with likeability, trust and satisfaction.

3.1.3.7 Previous work in trust measuring using psycho-physiological signals

There are few psycho-physiological measurements that have been studied in the context of human trust, as most of the experimentation like [44], [61]–[63] refer to questionnaires to evaluate human trust.

EEG is an electrophysiological measurement technique that captures the cortical activity of the brain. Some researchers have studied trust via EEG measurements, but only with event-related potentials (ERPs). ERPs measure brain activity in response to a specific event. An ERP is determined by averaging repeated EEG

[8] A system that enables a robot to robustly anticipate and adapt to other team members, make decisions on-the-fly, and consider the consequences of its actions on others

responses over many trials to eliminate random brain activity [64]. For instance, [65] found a difference in peak amplitudes of ERP components in human subjects while they participated in a coin toss experiment that stimulated trust and distrust. However, it is difficult to identify triggers during an actual human-machine interaction, thereby rendering ERPs impractical for real-time trust level sensing. On the other hand, some research like [66] found that several time-domain EEG features are significant to sense trust in humans.

GSR is a classical psycho-physiological signal that captures arousal based upon the conductivity of the surface of the skin. It is not under conscious control but is instead modulated by the sympathetic nervous system. GSR has also been used in measuring stress, anxiety, and cognitive load [67][68]. However, according to [69], the use of GSR for estimating trust has not been much explored and was noted as an area worth studying. [66] found that the net phasic component and the maximum value of phasic activity may be relevant in trust detection.

Pupillometry has also been recently introduced as a feasible measurement to detect human trust. For instance, [70] reveals that humans trust partners with dilating pupils and withheld trust from partners with constricting pupils. Even if no direct evidence is found regarding the use of pupillometry to detect trust in HRI, [71] shows that it is an indicator to estimate the high cognitive load states, whereas [72] points to an existing correlation between trust and cognitive load in HRC.

3.1.3.8 First conclusions

Previous sections key elements to conceptualize trust during HRC. As a summary, we may state that human body reacts to with certain psycho-physiological responses when exposed to different stimuli. These reactions, such as the heart rate, the brain activity or the galvanic skin response, create certain signals that can be measured and analysed in order to get a better understanding of the human nature. Unfortunately, most of the studies focus their results on the five basic Human Factors (valence, arousal, cognitive load, engagement and memory activation), which leads to unknowledge of how these psycho-physiological signals could be used to detect higher level cognitive processes such as trust.

On the other, the trusting process is presented as a complex response, affected with several controllable and uncontrollable factors. Therefore, it is necessary to identify those factors that may cause a major influence during Human-Robot relation and control them in order to improve trust towards a more collaborative environment between humans and robots.

Innovation in HR-Recycler about different Human Factors detected correlation between the classical Human Factors in Human-Robot Collaboration (valence, arousal, cognitive load, memory activation and engagement) and the concept of Trust of human in the robot. In this sense, to quantify the trust of human in robot, we will use the following Human Factors and signals:

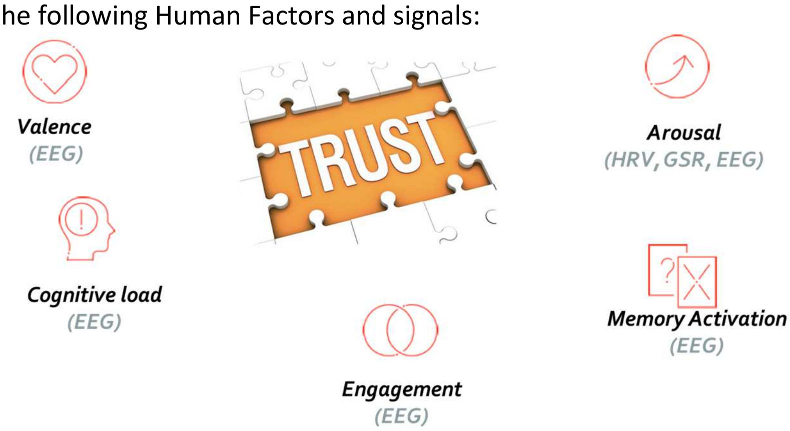


Figure 5: Human Factors related to Trust measurement

Also, considering the multiple factors that affect Human-Robot trust, we will determine the most relevant ones according both to their controllability and their impact in the construction and evolution of trust relationship.

3.2 Trust Factor Worker Model

In next release of this report, D7.2, we will provide a details Trust factor worker model based on the findings from the pilot studies performed and described in Section 4.

4 Pilot studies with end-users

In this section we define the specific pilot studies targeting end-users with a twofold purpose: first, understanding trust in collaborative disassembly and, second, validation of trust factor model to improve it.

As the pilot studies are difficult to perform in a real environment due to noise conditions and also to safety equipment that operators need to wear. However, the pilot studies will be done at two levels:

- In a Virtual Reality simulated environment.
- In a controlled environment, if possible, inside the plant.

We will use the following devices to measure the signals and human factors related to Trust:

Device	***Description***	***Image***
EEG – Electroencephalography	*Wearable and lightweight system for real-time EEG monitoring with maximum comfort and minimum intrusion. With a simple positioning, it allows registering with great reliability up to 12 dry EEG channels in optimized positions for the estimation of emotions and cognitive behaviours (pre-frontal, frontal, parietal, temporal and occipital cortex)*	
Biosensors – GSR (sweating) y BVP (pulse)	*Wearable and wireless team of bio-signal registration in real time to address all types of studies with ecology and quality. The device has an ultralight and comfortable design with two key biosensors for a basic estimation of emotions, and a three-axis accelerometer in solidarity for an estimation of the noise that can be generated due to the movement of the hand*	
Stationary Eye-Tracking	*Stationary eye tracking technology to measure the display patterns of stimuli that are displayed on the screen, either on PC or laptop*	
Versatile bio-signal amplifier	*Mobile and versatile bio-signal amplifier to monitor up to 35 physiological variables simultaneously with milli-second synchronization. It can record **up to 35 simultaneous channels** of analog bio-signals (GSR, ExG, RESP, TEMP, etc.), movement activity and location (EMG, IMUs, GPS, etc.)*	

Table 5: List of devices to measure the signals and human factors related to Trust

4.1 Pilot studies for understanding the Trust factor

The goal of the first pilot study is to understand the factors that cause variations in trust and identify the biological reactions related to it. These reactions will be validated as trust indicators in a further version of the experiment.

In order to expose participants to different trust stimuli towards machines, a variation of the Inspection Game have been designed. An inspection game is a mathematical model of a non-cooperative situation where an inspector verifies that another party, adheres to legal rules instead of shrinking work duties. However, in an HRC context there is no reason for a machine to elude the assigned job. Therefore, the new scenario is described as follows:

The participant is responsible of a manufacturing process with a machine to help him/her to perform the daily task. Due to external situations, the machine may be suffering some minor imbalance that affects the manufacturing process. If the machine is correctly set, the working station will produce 5 pieces per hour, but if the setting is imbalanced, only 1 piece will be produced hourly. The machine is equipped with a simple sensor that detects the balance status. In concordance with the sensor reading, which may be correct or not, the participant may decide to run an automatic checking process on the machine to warranty that it is correctly balanced, however, during the duration of this process the working station is inoperative. The duration of the checking process depends on the machine's real status. It will take 1 hour to complete if the machine is indeed unbalanced, but it will take 3 hours to complete if the machine is working correctly. The participant works on a 5-hour shift and (s)he must try to get as many pieces as possible.

Table 6 shows the payoff matrix of the experiment. It has been designed to be symmetrical to the participants' prior knowledge, that is, there is no initial reason to overrun one of the choices (check / No check) in favour of the other. In expected outcome in both cases is identical (15 pieces/hour) and, thus, this eliminates any kind of external influence in the analytical (trust / no trust) decision process. This table is provided to the participants and it is present at every moment during the experiment.

		Machine	
		Balanced	Unbalanced
Human	Check	10	20
	No check	25	5

Table 6: Payoff matrix of the designed experiment

Simplicity has been kept to a maximum during the design of this first experiment. Participants interact with a single screen that provides them the sensor feedback. The trust decision process is implemented via a single command bottom. Participants are exposed to 120 iterations during an approximate time of 30 minutes.

As the experiment runs, participants are exposed to different machine behaviours. First, they undertake a familiarization set, where the probability of the machine to fail and the probability of the sensor to wrongly detect such a fail are set either to 0% or to 100%, therefore granting a certainty regarding the machine behaviour. The purpose of this section is double: first, allow the participant to get familiar with the environment and to get to know the different messages that will rise during the experiment, and second, to record a prior set of responses where the participant if certain of the outcome of his/her actions in contrast which those to be recording during the experimental phase.

Once the familiarization finishes, they are exposed to the experiment phase of the test. During this test, the probability to rise a failure on the machine and the probability to wrongly detect this failure vary and are unknown to the participants. Thus, the psycho-physiological signals recorded during this phase are conditioned by the participants' uncertainty on the outcomes of their actions, and so, they are an indicator of the level of trust they set on the machine.

This experiment allows to detect how (un)predictability affects trust and how prior experiences affect trust. Both these factors are covered under the "learned trust" dimension described in section 3.1.3.5., which is the most trust dimension for the propose of the project since is the most controllable one.

Regarding the other two dimensions (dispositional and situational trust), the following considerations are taken into account:

- ***Dispositional trust*** is the most chaotic and less controllable aspect of trust, since it is inherent to the human nature of the participants. Furthermore, this dimension is heavily affected by the other dimension and, thus, its importance is limited to the very first interactions with the machines before any either the situational trust or the learned trust overpowers the natural inclinations. As the purpose of the project is based on long-term HRC, there is no specific protocol to capture this dimension of trust. However, during the data analysis process, the results will be segmented according to personal criteria such as age and/or gender. Considering the nature of this dimension of trust, it is likely that this segmentation process enables to discover some factors affecting this dimension.
- ***Situational trust*** plays an important role modulating the nature instincts reflected in the dispositional trust. If the whole scope of the project is considered, however, its relevance may decline since the collaborative relation between human and robot will always take place in the same or very similar circumstances (assembly tasks). Nevertheless, some of the factors described in section 3.1.3.5. are worthy to be considered in this analysis. For instance, the perceived task criticality may play a very significant role regarding the decision of trust. In order to check the importance of this factor, slight variations of the trust game will be performed to different set of individuals, altering to payoff matrix in order to increase the gap between adequate or inadequate decisions or creating an extraordinary stress factor (for instance, setting an objective amount of pieces to be manufactured in a certain number of interactions). The variations of robot behaviour will also be analysed, but on a second iteration of the process (during the validation experiences) by means of virtual reality approaches.

So, considering the previous facts, dispositional trust is not really a controllable dimension of trust, and so, its influence will be considered in the study using data segmentation, but we won't be setting any experimental feature to detect the influence of this factor in within the global trust. On the other hand, there are features that condition situational trust and could be used to modulate human trust, so they are worthy to be analysed. However, as the first experiment will be oriented to detect which psycho-physiological signals may be more correlated to trust/distrust situations, the inclusion of further variation may difficult the experimental procedure and, thus, these factors will be included and analysed on a second experimental phase which will be conducted on a VR environment.

4.2 Pilot studies for validation of the Trust factor model

In next version of this deliverable more details will be provided about these pilot studies.

5 Assessment of robot social behaviours

In this section, we describe the assessment of the interactive capabilities of the HR-Recycler robots. The goal is to develop robots that will not only be accepted by humans but also promote and improve collaboration between them. More specifically, we aim to evaluate a series of robot behaviours that include effective and timely communication channels as well as robot behaviours (such as gestures and feedback). The outcomes of these pilot studies will inform the worker model (T4.2), the integration of social cognition and adaptive behaviours to robots (T7.2). All studies will comply with the standards indicated by WP2 "Regulatory, legal, ethical and societal challenges of robotics in industrial automation" and the requirements of local Ethical Committees.

First, we will aim to identify the necessary requirements for an effective communication and collaboration between non-humanoid robots and human co-workers. As mentioned in section 2, communication is important for successful collaboration and a variety of channels can be employed. In the context of HR-Recycler, we have identified two main channels of communication: gestures and interactive mediums such as a tablet or Augmented Reality.

Robot gestures and behaviour have been extensively studied in HRC; however, little attention has been made when it comes to industrial settings where robots are non-humanoid. This poses a challenge to the correct interpretation of the message and intention conveyed by gestures. Indeed, studies suggest that successful collaboration relies on correct interpretation of gestures [76]. Additionally, successful recognition is likely to determine the human co-worker trust. For example, robot behaviour and predictability highly influence trust [77]. Trust has been extensively studied in HRC, but only few studies evaluate the impact of non-humanoid robots on human trust. As a first step, we will devise a set of suitable gestures for the HR-Recycler's use-cases that we will evaluate by asking participants to report on what these gestures mean to communicate. At a follow-up study, the gestures that scored higher in recognition (or are relevant for the task) will be used in a task where humans need to collaborate with a robot. Here, the robot will employ these gestures to communicate for the successful completion of the task. In this study, we will measure task performance, subjective trust and self-confidence as well as overall robot and task evaluation by using standardised questionnaires followed by a short semi-structured interview. The outcomes of this study will inform the social behaviour of the robot (T7.2).

Although gestures are important for collaboration, they cannot cover the full range of possible interactions, especially when complex messages need to be communicated such as task instructions or even alerts whose content may vary according to the situation. Given the noisy nature of the environment, employing high-level dialogue is not possible. For this reason, we have identified two channels that have been employed in industrial settings: Augmented Reality or handheld devices (such as tablets). Here, participants will interact with a non-humanoid robot in a collaborative task through either AR or a tablet. To understand which of the presentation means is more effective, we will assess task performance and robot evaluation using standardised questionnaires among other recorded data.

Although using the appropriate communication channel is important for successful collaboration, a critical component that has not been extensively studied is when to communicate. In most cases, the human co-worker will be focused in the task at hand. Depending on the situation and the criticality of the message, the robot can choose to either interrupt the co-worker or wait until the co-worker finishes the task. In this pilot study we will evaluate under which conditions interrupting (or not) the co-worker is preferred and leads to successful task completion as well as how this affects human trust. The data collected will be similar to the evaluation of robot gestures. Outcomes of this study will not only inform the social behaviour of the robot (T7.2) but also the Worker Model (T4.2), as different users may have different preferences.

Finally, we plan to evaluate the feedback provided by the robot as well as the complexity of the interaction. Studies suggest that sociality is an important factor in collaboration, even in industrial settings [78]. Human co-workers seem to desire social interactions with robots and expect them to display basic conversational skills such as greetings at the start or end of a shift. To facilitate this coordination, we will evaluate the communicative functions and robot behaviour within the context of a collaborative task and assess how such social behaviours affect collaboration, trust and perceived safety. The outcomes of this study will inform the social behaviour of the robot (T7.2) and the worker model (T4.2).

By designing collaborative robots in a human-centred approach, we can understand the dynamics in the interactions between humans and robots and promote positive collaborations. The proposed studies aim at developing and evaluating multimodal communication frameworks that could support more complex interactions and behaviours, a domain that is in its infancy and efficient methods of interaction are still lacking [79].

6 Conclusions and future work

In HR-Recycler, HRC is based on supervisory control of robots' performance by humans in routine tasks (e.g. disassembling and sorting electronical devices and components). Considering that Human-Robot Collaboration triggers some emotions in the human, the goal of this deliverable is to quantify and detect these emotions and responses in order to design a smoother collaboration and interaction among humans and robots. For doing so, this report provides an initial definition of communication requirements, putting special emphasis on advancing on mechanisms that underlie social competence in a broader range of non-human social behaviours for communication and collaboration. Besides, in next version of this deliverable novel and intuitive ways of delivering natural and transparent communication and interaction will be explored.

After a thorough analysis of the state of the art in HRI and HRC, **Trust** has been identified a major factor to determine effective collaboration among agents, both human-to-human and human-to-machine, alike. We distinguish between modelling and measuring trust modelling aims to represent the nature of human trust behaviour extrapolating individual responses to a universal level, while trust measuring looks for mensurable involuntary body responses when the subjects are exposed to different trust-related stimuli. Based on this, we have defined a set of specific pilot studies targeting end-users to understand the trust in HRC. Based on the main findings from SotA and from these pilot studies, the next release of this report, D7.2, will provide a details Trust factor worker model that we will also validate at two levels: (1) in a Virtual Reality simulated environment and (2) in a controlled environment, if possible, inside the plant.

Finally, we propose a set of pilot studies that aim to assess the interactive capabilities of the robots, so they are accepted by humans and foster collaboration. We aim to systematically evaluate the necessary requirements for effective communication and collaboration as well as understand the dynamics in interactions between humans and robots. We propose to evaluate the robot's communication channels and behaviours in a series of studies that will inform the worker model and the social cognition and adaptive behaviour of the robot (tasks 4.2 and 7.2 respectively).

7 References

[1] Öhman A, Hamm A, Hugdahl K. Cognition and the autonomic nervous system: Orienting, anticipation, and conditioning. In: Cacioppo JT, Tassinary LG, Berntson GG, editors. Handbook of psychophysiology. 2. New York: Cambridge University Press; 2000. pp. 533–575.

[2] Levenson, R. W. (2014). The autonomic nervous system and emotion. Emotion Review, 6(2), 100–112. doi:10.1177/1754073913512003.

[3] Bekey, G., & Yuh, J. (2008). The status of robotics. Robotics & Automation Magazine, IEEE, 15(1), 80–86. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4476332

[4] Bar-Cohen, Y., & Hanson, D. (2009). The coming robot revolution: Expectations and fears about emerging intelligent, humanlike machines. Retrieved from http://books.google.com/books?hl=fr&lr=&id=S0KJg6pW14QC&oi=fnd&pg=PR5&dq=The+Coming+Robot+Revolution:+Expectations+and+Fears+About+Emerging+Intelligent&ots=IceM86tkLJ&sig=VKSlpGs_PEA9k2p9vS0xWHbg3CI

[5] Sung, J. (2009). Robots in the wild: understanding long-term use. Human-Robot Interaction (…, 45–52. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6256092

[6] Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. Robotics and Autonomous Systems. Retrieved from http://www.sciencedirect.com/science/article/pii/S092188900200372X

[7] Goodrich, M. A., & Schultz, A. C. (2008). Human–robot interaction: a survey. Foundations and Trends® in Human–Computer Interaction, 1(3), 203–275. https://doi.org/10.1561/1100000005

[8] Leite, I., Martinho, C., & Paiva, A. (2013). Social Robots for Long-Term Interaction: A Survey. International Journal of Social Robotics, 5(2), 291–308. https://doi.org/10.1007/s12369-013-0178-y

[9] Reeves, B., & Nass, C. (1996). The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places (CSLI Lecture Notes S), (August 2016), 305. https://doi.org/10.1109/MSPEC.1997.576013

[10] Knapp, M. L., Hall, J. A., & Horgan, T. G. (2013). Nonverbal Communication in Human Interaction, 530. https://doi.org/10.1080/17404620601014724

[11]Johal, W., Calvary, G., & Pesty, S. (2014). Non-verbal Signals in HRI: Interference in Human Perception. 6th International Conference, ICSR 2014, Proceedings, 412. https://doi.org/10.1007/978-3-319-11973-1

[12]Mattessich, P. W., & Monsey, B. R. (1992). Collaboration: what makes it work. A review of research literature on factors influencing successful collaboration.

[13]Cohen, P. R., & Levesque, H. J. (2014). Teamwork. Nous, 25(4), 487–512.

[14] Grosz, B. (1996). Collaborative Systems (AAAI-94 Presidential Address). AI Magazine, 17(2), 67–85. https://doi.org/10.1609/aimag.v17i2.1223

[15] Levinson, S. (2006). Cognition at the heart of human interaction. Discourse Studies, 8, 85–93. https://doi.org/10.1177/1461445606059557

[16]Michalos, G., Karagiannis, P., Makris, S., Tokçalar, Ö., & Chryssolouris, G. (2016). Augmented Reality (AR) Applications for Supporting Human-robot Interactive Cooperation. Procedia CIRP, 41, 370–375. https://doi.org/10.1016/j.procir.2015.12.005

[17] Ong, S. K., Pang, Y., & Nee, A. Y. C. (2007). Augmented reality aided assembly design and planning. CIRP Annals - Manufacturing Technology, 56(1), 49–52. https://doi.org/10.1016/j.cirp.2007.05.014

[18] Rentzos, L., Papanastasiou, S., Papakostas, N., & Chryssolouris, G. (2013). Augmented reality for human-based assembly: Using product and process semantics. IFAC Proceedings Volumes (IFAC-PapersOnline) (Vol. 12). IFAC. https://doi.org/10.3182/20130811-5-US-2037.00053

[19] Fang, H. C., Ong, S. K., & Nee, A. Y. C. (2012). Interactive robot trajectory planning and simulation using augmented reality. Robotics and Computer-Integrated Manufacturing, 28(2), 227–237. https://doi.org/10.1016/j.rcim.2011.09.003

[20] Cassell, J., Bickmore, T., Vilhjhlmsson, H., & Yan, H. (2000). More Than Just a Pretty Face : Affordances of Embodiment. In Proceedings of the 5th international conference on Intelligent user interfaces (pp. 52–59).

[21] Boyle, E. A., Anderson, A. H., & Newlands, A. (1994). The Effects of Visibility on Dialogue and Performance in a Cooperative Problem Solving Task. Language and Speech, 37(1), 1–20. https://doi.org/10.1177/002383099403700101

[22] Bainbridge, W. a., Hart, J., Kim, E. S., & Scassellati, B. (2008). The effect of presence on human-robot interaction. RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication, 701–706. https://doi.org/10.1109/ROMAN.2008.4600749

[23] Powers, A., Kiesler, S., Fussell, S., & Torrey, C. (2007). Comparing a Computer Agent with a Humanoid Robot. ACM/IEEE International Conference on Human-Robot Interaction (HRI), (January), 145–152. https://doi.org/10.1145/1228716.1228736

[24]Goetz, J., Kiesler, S., & Powers, a. (2003). Matching robot appearance and behaviour to tasks to improve human-robot cooperation. The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003., 55–60. https://doi.org/10.1109/ROMAN.2003.1251796

[25] Walters, M. L., Syrdal, D. S., Dautenhahn, K., Te Boekhorst, R., & Koay, K. L. (2008). Avoiding the uncanny valley: Robot appearance, personality and consistency of behaviour in an attention-seeking home scenario for a robot companion. Autonomous Robots, 24(2), 159–178. https://doi.org/10.1007/s10514-007-9058-3

[26] Hinds, P., Roberts, T., & Jones, H. (2004). Whose Job Is It Anyway? A Study of Human-Robot Interaction in a Collaborative Task. Human-Computer Interaction, 19, 151–181. https://doi.org/10.1207/s15327051hci1901&2_7

[27] Bauer, Andrea & Wollherr, Dirk & Buss, Martin. (2008). Human-Robot Collaboration: a Survey. I. J. Humanoid Robotics. 5. 47-66. 10.1142/S0219843608001303.

[28] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and T. P. Group, "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," PLOS Med., vol. 6, no. 7, pp. 1–6, 2009.

[29] Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Eds.), Cognitive neuroscience of emotion (pp. 242 –276). New York, NY: Oxford University Press

[30] Russell, James A. Core affect and the psychological construction of emotion. Psychological Review, Vol 110(1), Jan 2003, 145-172

[31] Russell JA, Barrett LF. 1999. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. Journal of Personality and Social Psychology 76(5):805819

[32] Posner J, Russell JA, Peterson BS. 2005. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and Psychopathology 17(3):715734

[33] Schaufeli, W.B. (2013). What is engagement? In C. Truss, K. Alfes, R. Delbridge, A. Shantz, & E. Soane (Eds.), Employee Engagement in Theory and Practice. London: Routledge.

[34] Stuart Ira Fox. Human Physiology. Mcgraw-Hill (Tx), 7 edition, 2002.

[35]P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm, "Looking at pictures: affective, facial, visceral, and behavioural reactions," Psychophysiology, vol. 30, pp. 261-273, 199.

[36] J.-H. Cho, K. Chan, and S. Adali, 'A Survey on Trust Modeling', ACM Comput. Surv., vol. 48, no. 2, pp. 28:1–28:40, Oct. 2015.

[37] D. Gambetta, 'Can we trust', Trust: Making and breaking cooperative relations, vol. 13, pp. 213–237, 2000.

[38] B. Lahno, 'Olli lagerspetz: Trust. The tacit demand', Ethical Theory and Moral Practice, vol. 2, no. 4, pp. 433–435, 1999.

[39] H. S. James Jr, 'The trust paradox: a survey of economic inquiries into the nature of trust and trustworthiness', Journal of Economic Behavior & Organization, vol. 47, no. 3, pp. 291–307, 2002.

[40] J. B. Rotter, 'Interpersonal trust, trustworthiness, and gullibility.', American psychologist, vol. 35, no. 1, p. 1, 1980.

[41] F. D. Schoorman, R. C. Mayer, and J. H. Davis, 'An integrative model of organizational trust: Past, present, and future', 2007.

[42] A. H. Kydd, Trust and mistrust in international relations. Princeton University Press, 2007.

[43] J. D. Lee and K. A. See, 'Trust in automation: Designing for appropriate reliance', Human factors, vol. 46, no. 1, pp. 50–80, 2004.

[44] J.-H. Cho, A. Swami, and R. Chen, 'A survey on trust management for mobile ad hoc networks', IEEE Communications Surveys & Tutorials, vol. 13, no. 4, pp. 562–583, 2010.

[45] E. M. Uslaner, The moral foundations of trust. Cambridge University Press, 2002.

[46] K. A. Hoff and M. Bashir, 'Trust in automation: Integrating empirical evidence on factors that influence trust', Human factors, vol. 57, no. 3, pp. 407–434, 2015.

[47] D. M. Romano, 'The nature of trust: conceptual and operational clarification', 2003.

[48] K. Schaefer, 'The perception and measurement of human-robot trust', 2013.

[49] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, 'A meta-analysis of factors affecting trust in human-robot interaction', Human factors, vol. 53, no. 5, pp. 517–527, 2011.

[50] K. Dautenhahn, 'Socially intelligent robots: dimensions of human–robot interaction', Philosophical transactions of the royal society B: Biological sciences, vol. 362, no. 1480, pp. 679–704, 2007.

[51] K. Dautenhahn, 'Robots we like to live with?!-a developmental perspective on a personalized, life-long robot companion', presented at the RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759), 2004, pp. 17–22.

[52] B. Sadrfaridpour, H. Saeidi, J. Burke, K. Madathil, and Y. Wang, 'Modeling and control of trust in human-robot collaborative manufacturing', in Robust Intelligence and Trust in Autonomous Systems, Springer, 2016, pp. 115–141.

[53] D. A. Wiegmann, A. Rich, and H. Zhang, 'Automated diagnostic aids: The effects of aid reliability on users' trust and reliance', Theoretical Issues in Ergonomics Science, vol. 2, no. 4, pp. 352–367, Jan. 2001.

[54] N. Moray, T. Inagaki, and M. Itoh, 'Adaptive automation, trust, and self-confidence in fault management of time-critical tasks.', Journal of experimental psychology: Applied, vol. 6, no. 1, p. 44, 2000.

[55] W. A. Bainbridge, J. Hart, E. S. Kim, and B. Scassellati, 'The effect of presence on human-robot interaction', presented at the RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication, 2008, pp. 701–706.

[56] J. Shah, J. Wiken, B. Williams, and C. Breazeal, 'Improved human-robot team performance using Chaski, A human-inspired plan execution system', in 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2011, pp. 29–36.

[57] E. J. de Visser et al., 'Almost human: Anthropomorphism increases trust resilience in cognitive agents.', Journal of Experimental Psychology: Applied, vol. 22, no. 3, p. 331, 2016.

[58] K. Akash, T. Reid, and N. Jain, 'Improving Human-Machine Collaboration Through Transparency-based Feedback–Part II: Control Design and Synthesis', IFAC-PapersOnLine, vol. 51, no. 34, pp. 322–328, 2019.

[59] K. Fischer, H. M. Weigelin, and L. Bodenhagen, 'Increasing trust in human–robot medical interactions: effects of transparency and adaptability', Paladyn, Journal of Behavioral Robotics, vol. 9, no. 1, pp. 95–109, 2018.

[60] D. Li, P. P. Rau, and Y. Li, 'A cross-cultural study: Effect of robot appearance and task', International Journal of Social Robotics, vol. 2, no. 2, pp. 175–186, 2010.

[61] S. Soroka, J. F. Helliwell, and R. Johnston, 'Measuring and modelling trust', Diversity, social capital and the welfare state, pp. 279–303, 2003.

[62] K. Leichtenstern, N. Bee, E. André, U. Berkmüller, and J. Wagner, 'Physiological Measurement of Trust-Related Behavior in Trust-Neutral and Trust-Critical Situations', in Trust Management V, Berlin, Heidelberg, 2011, pp. 165–172.

[63] T. Nomura and S. Takagi, 'Exploring effects of educational backgrounds and gender in human-robot interaction', in 2011 International Conference on User Science and Engineering (i-USEr), 2011, pp. 24–29.

[64] T. C. Handy, Event-related potentials: A methods handbook. MIT press, 2005.

[65] C. Boudreau, M. D. McCubbins, and S. Coulson, 'Knowing when to trust others: An ERP study of decision making after receiving information from unknown people', Social cognitive and affective neuroscience, vol. 4, no. 1, pp. 23–34, 2008.

[66] K. Akash, W.-L. Hu, N. Jain, and T. Reid, 'A classification model for sensing human trust in machines using eeg and gsr', ACM Transactions on Interactive Intelligent Systems (TiiS), vol. 8, no. 4, p. 27, 2018.

[67] S. C. Jacobs et al., 'Use of skin conductance changes during mental stress testing as an index of autonomic arousal in cardiovascular research', American heart journal, vol. 128, no. 6, pp. 1170–1177, 1994.

[68] R. Nikula, 'Psychological correlates of nonspecific skin conductance responses', Psychophysiology, vol. 28, no. 1, pp. 86–90, 1991.

[69] R. Riedl and A. Javor, 'The biology of trust: Integrating evidence from genetics, endocrinology, and functional brain imaging.', Journal of Neuroscience, Psychology, and Economics, vol. 5, no. 2, p. 63, 2012.

[70] M. Kret, A. Fischer, and C. K. De Dreu, 'Pupil mimicry correlates with trust in in-group partners with dilating pupils', Psychological science, vol. 26, no. 9, pp. 1401–1410, 2015.

[71] G. Minadakis and K. Lohan, 'Using Pupil Diameter to Measure Cognitive Load', arXiv preprint arXiv:1812.07653, 2018.

[72] M. I. Ahmad, J. Bernotat, K. Lohan, and F. Eyssel, 'Trust and Cognitive Load During Human-Robot Interaction', arXiv preprint arXiv:1909.05160, 2019.

[73] B. Gonsior, M. Buß, S. Sosnowski, D. Wollherr, K. Kuhnlenz, and M. Buss, 'Towards transferability of theories on prosocial behavior from Social Psychology to HRI', Proceedings of IEEE Workshop on Advanced Robotics and its Social Impacts, ARSO, p. 101-103, 2012.

[74] B. Kühnlenz, S. Sosnowski, M. Buß, D. Wollherr, K. Kühnlenz, and M. Buss, 'Increasing Helpfulness towards a Robot by Emotional Adaption to the User', International Journal of Social Robotics, vol. 5, no. 4, p. 457-476, 2013.

[75] A. Turnwald and D. Wollherr, 'Human-Like Motion Planning Based on Game Theoretic Decision Making', International Journal of Social Robotics, vol. 11, no. 1, p. 151-170, 2018.

[76] B. Gleeson, K. Maclean, A. Haddadi, E. Croft, and J. Alcazar, 'Gestures for industry: Intuitive human-robot communication from human observation', ACM/IEEE International Conference on Human-Robot Interaction, p. 349-356, 2013.

[77] Sanders, T., Oleson, K. E., Billings, D. R., Chen, J. Y., & Hancock, P. A. (2011, September). 'A model of human-robot trust: Theoretical model development', in Proceedings of the human factors and ergonomics society annual meeting (Vol. 55, No. 1, pp. 1432-1436). Sage CA: Los Angeles, CA: SAGE Publications.

[78] Sauppé, A. and Mutlu, B., 'The social impact of a robot co-worker in industrial settings', in Proceedings of the 33rd annual ACM conference on human factors in computing systems pp. 3613-3622, 2015.

[79] Tsarouchi, P., Makris, S. and Chryssolouris, G., 'Human–robot interaction review and challenges on task planning and programming', International Journal of Computer Integrated Manufacturing, 29(8), pp.916-931, 2016.

[80] Ahmad Khawaji, Jianlong Zhou, Fang Chen, and Nadine Marcus. 2015. Using Galvanic Skin Response (GSR) to Measure Trust and Cognitive Load in the Text–Chat Environment. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. ACM Press, 1989–1994.

[81] Reiner Nikula. 1991. Psychological Correlates of Nonspecific Skin Conductance Responses. Psychophysiology 28, 1. (1991), 86–90.

[82] Ahmad Khawaji, Jianlong Zhou, Fang Chen, and Nadine Marcus. 2015. Using Galvanic Skin Response (GSR) to Measure Trust and Cognitive Load in the Text–Chat Environment. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. ACM Press, 1989–1994.

[83] René Riedl and Andrija Javor. 2012. The biology of trust: Integrating evidence from genetics, endocrinology, and functional brain imaging. Journal of Neuroscience, Psychology, and Economics 5, 2 (2012), 63.

[84] Vuilleumier, P., & Sander, D. (2008). Trust and valence processing in the amygdala. Social cognitive and affective neuroscience, 3(4), 299-302.

[85] Schniter, E., Sheremeta, R. M., & Shields, T. W. (2015). Conflicted emotions following trust-based interaction. Journal of Economic Psychology, 51, 48-65.

[86] Nikolaidis, S., Hsu, D., & Srinivasa, S. (2017). Human-robot mutual adaptation in collaborative tasks: Models and experiments. The International Journal of Robotics Research, 36(5-7), 618-634.

[87] Gombolay, M. C., Gutierrez, R. A., Clarke, S. G., Sturla, G. F., & Shah, J. A. (2015). Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams. Autonomous Robots, 39(3), 293-312.
[88] Shah, J., Wiken, J., Williams, B., & Breazeal, C. (2011). Improved human-robot team performance using chaski, a human-inspired plan execution system. In Proceedings of the 6th international conference on Human-robot interaction (pp. 29-36).
[89] Zhang, J., Wang, Y., & Xiong, R. (2016, August). Industrial robot programming by demonstration. In 2016 International Conference on Advanced Robotics and Mechatronics (ICARM) (pp. 300-305). IEEE.
[90] Lee, J. (2017). A survey of robot learning from demonstrations for human-robot collaboration. arXiv preprint arXiv:1710.08789.
[91] Moulin-Frier, C., Fischer, T., Petit, M., Pointeau, G., Puigbo, J.-Y., Pattacini, U., … Verschure, P. F. M. J. (2017). DAC-h3: A Proactive Robot Cognitive Architecture to Acquire and Express Knowledge About the World and the Self. IEEE Transactions on Cognitive and Developmental Systems, 1–1. doi:10.1109/tcds.2017.2754143
[92] Misra, I., Girshick, R., Fergus, R., Hebert, M., Gupta, A., & Van Der Maaten, L. (2018). Learning by asking questions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 11-20).
[93] Cakmak, M., & Thomaz, A. L. (2012, March). Designing robot learners that ask good questions. In 2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 17-24). IEEE.
[94] Wang, W., Li, R., Diekel, Z. M., & Jia, Y. (2018). Robot action planning by online optimization in human–robot collaborative tasks. International Journal of Intelligent Robotics and Applications, 2(2), 161–179. doi:10.1007/s41315-018-0054-x
[95] Kattepur, A., & Balamuralidhar, P. (2019). RoboPlanner: Towards an Autonomous Robotic Action Planning Framework for Industry 4.0.
[96] Balakirsky, S. (2015). Ontology based action planning and verification for agile manufacturing. Robotics and Computer-Integrated Manufacturing, 33, 21–28. doi:10.1016/j.rcim.2014.08.011

ANNEX 1 – Summary of Literature Review

Studies	***Defined Criteria***		***Global Rating***
	HF & EM	**HF & HRC**	
Öhman A, Hamm A, Hugdahl K. (2000). Cognition and the autonomic nervous system: Orienting, anticipation, and conditioning. In: Cacioppo JT, Tassinary LG, Berntson GG, editors. Handbook of psychophysiology. 2. New York: Cambridge University Press; pp. 533–575.	M	W	M
Levenson, R. W. (2014). The autonomic nervous system and emotion. Emotion Review, 6(2), 100–112. doi:10.1177/1754073913512003.	S	W	M
Bauer, Andrea & Wollherr, Dirk & Buss, Martin. (2008). Human-Robot Collaboration: a Survey. I. J. Humanoid Robotics. 5. 47-66. 10.1142/S0219843608001303.	M	S	S
D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and T. P. Group,(2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement, PLOS Med., vol. 6, no. 7, pp. 1–6.	W	W	W
Stuart Ira Fox (2002). Human Physiology. Mcgraw-Hill (Tx), 7 edition.	S	W	M
P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm (1999). Looking at pictures: affective, facial, visceral, and behavioural reactions, Psychophysiology, vol. 30, pp. 261-273.	S	M	S
Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Eds.), Cognitive neuroscience of emotion (pp. 242 –276). New York, NY: Oxford University Press	M	W	M
James A. (2003). Core affect and the psychological construction of emotion. Russell, Psychological Review, Vol 110(1), 145-172.	M	W	W
Schaufeli, W.B. (2013). What is engagement? In C. Truss, K. Alfes, R. Delbridge, A. Shantz, & E. Soane (Eds.), Employee Engagement in Theory and Practice. London: Routledge	M	W	W
Russell JA, Barrett LF. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. Journal of Personality and Social Psychology 76(5):805819	M	W	W
Posner J, Russell JA, Peterson BS. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and Psychopathology 17(3):715734	M	W	M

Table 7: Quality assessment (W: Weak, M: Moderate, S: Strong) of included studies reporting Human Factors based on the defined criteria (HF & EM: Human Factors and emotion model and HF & HRC: more specific for Human Factors and Human-Robot Collaboration).