Expectable Motion Unit: Avoiding Hazards From Human Involuntary Motions in Human-Robot Interaction

Robin Jeanne Kirschner, Henning Mayer, Lisa Burr, Nico Mansfeld, Saeed Abdolshah, and Sami Haddadin

***Abstract*— In robotics, many control and planning schemes have been developed that ensure the human physical safety in human-robot interaction. The human psychological state and expectation towards the robot, however, are typically neglected. Even if the robot behaviour is regarded as biomechanically safe, humans may still react with rapid involuntary motion (IM) caused by startle or surprise. Obviously, such sudden, uncontrolled motions can jeopardize safety and should be prevented by any means. In this paper, we propose the Expectable Motion Unit (EMU) concept which ensures that a certain probability of IM occurrence is not exceeded in a typical HRI setting. Based on a model of IM occurrence that we generate through an experiment with 29 participants, the mapping between robot velocity, robot-human distance, and the relative frequency of IM occurrence is established. This mapping is processed towards a real-time capable robot motion generator, which limits the robot velocity during task execution if necessary. The EMU is combined with the well-established Safe Motion Unit in order to integrate both physical and psychological safety knowledge and data into a holistic safety framework. In a validation experiment, it was shown that the EMU successfully avoids human IM in five out of six cases.**

I. INTRODUCTION

Safety is a key requirement for the successful implementation of modern collaborative robots in real-world industrial and service scenarios. As proximity is an essential part of smooth human-robot interaction, collisions and contact (desired, undesired, or even unforeseen) may occur. In robotics, many pre- and post-collision strategies have been introduced to ensure the human physical integrity, e.g., collision detection and reaction [1], collision avoidance [2], [3], and real time model-, metrics-, or injury data-based control [4], [5]. Besides ensuring the human's physical integrity, safe and efficient human-robot interaction (HRI) also requires understanding and estimation of the human state, behaviour, and responses, e.g., human pose estimation [6], [7], [8] or affection towards robots [9]. An important factor that should be considered in HRI is the human expectation [10]. If the expectation is violated, then the human can react with startle and surprise [11]. This includes rapid involuntary human motions (IM), which may jeopardize safety [12].

To avoid possibly hazardous contacts even in case of IM, several authors assume the worst case human motion range and dynamics in their control and planning schemes [3], [2]. Such schemes can become overly conservative and may lead to large separation distances between human and robot, even if the probability of IM is low. Safer, closer, and more time-efficient HRI can be achieved if human IM occurrence (IMO) is reduced or excluded.

Robin Jeanne Kirschner, Lisa Burr, Nico Mansfeld, Saeed Abdolshah, and Sami Haddadin are with the Munich Institute of Robotics and Machine Intelligence (MIRMI), formerly MSRM, Technical University of Munich, Munich, Germany. Henning Mayer is with Munich Center for Technology and Society, Technical University of Munich, Munich, Germany, robin-jeanne.kirschner@tum.de.

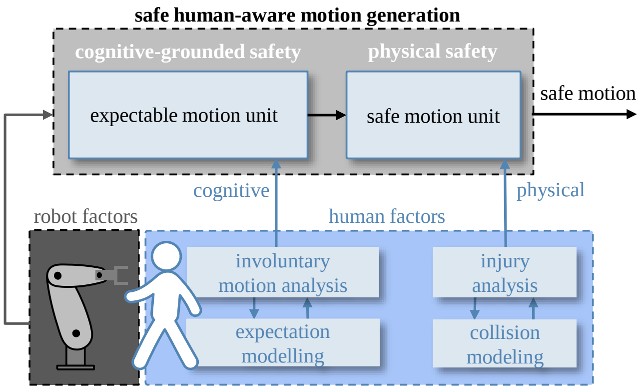

Fig. 1. Framework for safe human-aware motion generation combining the cognitive-grounded safety criterion involuntary motion occurrence and the well-established physical safety criterion injury occurrence. Experimentally established models for human injury and the human expectation towards robot motion are applied to generate safe robot motion such that it converges to the human expectation reducing involuntary motion occurrence.

In this paper, we develop a systematic approach to improve the performance and safety in HRI by avoiding human IM. First, we investigate the influence of robot motion parameters on the probability of human IMO in a common use case via an exploratory study involving 29 participants. The collected data and knowledge are then processed towards a human-aware, real-time capable motion generator that limits the robot speed so that a certain IMO probability is not exceeded. This safety tool is called the *Expectable Motion Unit* (EMU). It can seamlessly be combined with state-of-the-art safety schemes for avoiding human injury or pain during collisions, such as the Safe Motion Unit (SMU) which was proposed in [13]; see Fig. 1. A preliminary validation experiment involving eleven participants finally shows that the proposed framework successfully avoids for IM in the considered use case. Overall, the EMU concept aims at improving human safety, human-robot team performance, and robot user trust enabling safe and trustworthy HRI.

The remainder of the paper is organized as follows. In Sec. II we give a brief overview of the related work. In Sec. III we introduce our general approach for EMU aiming at preventing IMO. We conduct an exploratory experiment and evaluation method to derive the risk of IMO and represent it in a risk matrix in Sec. IV. A description of the implementation of the EMU velocity scaling is presented in Sec. V followed by an validation experiment. Sec. VI discusses the results and

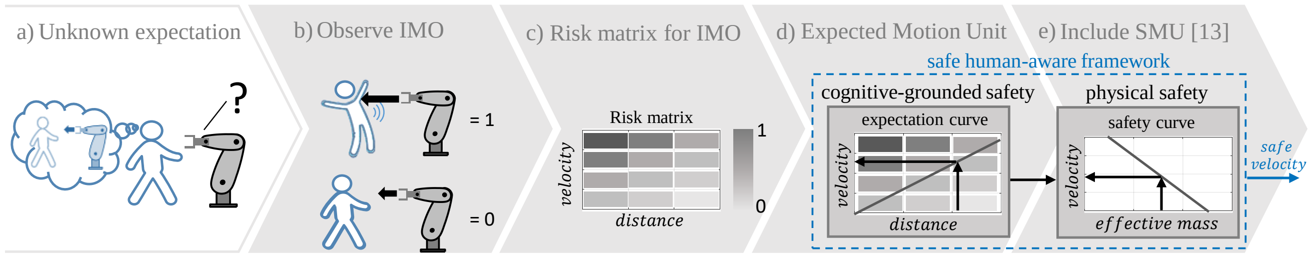

Fig. 2. EMU approach from the unknown human expectation to safe velocities based on human physical safety requirements and the reduction of IMO by experimental observation and modelling as risk matrix and expectation curve.

future direction of research. Finally, Sec. VII concludes the paper.

II. Related Work

In this section, we briefly summarise the considered physical safety approach and work on the human perception of safety in close HRI, in particular the influence of robot velocity and human-robot distance.

A central goal in HRI is to ensure that no injury occurs in any collision scenario between robot and human [14]. In this paper, we consider the Safe Motion Unit (SMU) proposed in [13] for safe robot velocity scaling to avoid human injury in HRI. The robot effective mass at the contact location in the Cartesian unit direction of motion, in other words, the perceived mass during contact [15], and robot velocity are mapped to the human injury at a collision. The human injury data is associated to certain human body parts, injury levels, and impact surface geometry at the contact location. Then, safe velocities are defined for a collision scenario via a so-called safety curve.

Embedding psychological aspects, e.g., human comfort to robot motion control is suspect to various research [16], [17], [18]. Instead of increasing comfort, we want to examine the circumstances in which - roughly speaking - people experience discomfort, and therefore respond with IM. Thus, we focus on the basic concepts for understanding potential influencing factors to IMO. IM as well as other cognitive impairments results from startle reflexes and surprise reactions which are potentiated by fear [12]. Literature on how startle and surprise can be identified and how they relate to each other is diverse and requires context dependent analysis. According to [19], both startle and surprise lie on a continuum where startle is a strong surprise reaction. Thus, both types of responses can negatively impact safety in HRI environments and can be observed through multimodal social signal coding. Social cues, which follow unexpected robot movements are, e.g., felt smiles [20] or body freezes [21]. The human expectation which shapes the reaction towards a robot motion in a certain event is influenced by the human perception of the scenario [11].

Research in HRI especially considers the perception of human safety [22] and influencing factors. One parameter influencing the perception of safety in HRI is the instantaneous human-robot distance which can be explained by proxemic behaviour models [23]. The model of proxemics

TABLE I
Robot Acceptance Social Zones regarding "Proxemics"

Personal space zone	Range	Supporting studies
Close intimate	0.00 m - 0.15 m	[25]
Intimate	0.15 m - 0.45 m	[25], [26]
Personal	0.45 m - 1.20 m	[25], [26], [27]
Social	1.20 m - 3.60 m	
Public	$\geq$ 3.60 m	

is introduced in [24] and describes different social zones in which humans like to interact with each other. It is used in studies on mobile robots, based on the assumption that robots are treated as social instances [25], [26], [27]. Unexpectedly, in studies with the mechanical looking *PeopleBot*, it was found that $40\ \%$ of the participants were comfortable with the robot in proximity of less than $0.45\,\mathrm{m}$ where usually only intimate human relationships are accepted. This suggests that not everyone regards the robot as a social instance and even closer human-robot proximity is still acceptable[25]. Tab. I lists which human-robot distances were accepted by the participants of [25], [26], [27] in relation to the well-established proxemics model. In addition to the distance between human and robot, several studies conclude that robot velocity is a factor that strongly correlates with the feeling of arousal [28], perceived level of hazard [29], and perceived safety [30]. Other studies also investigated robot approaching motion parameters including varying motion direction, acceleration, and jerk in the context of human-robot handover scenarios [31].

Based on the presented findings of related research we propose our concept for involuntary motion avoidance in the following.

III. Expectable Motion Unit Concept

In this work, we propose a cognitive-grounded safety concept based on the human expectation fulfillment the so-called Expectable Motion Unit approach. The EMU aims to ensure a robot performs motions which are expected by the human and thus avoids human involuntary motions in HRI by velocity scaling based on a model of human IMO; see Fig. 1. The EMU approach improves safety as well as performance and can be combined with control schemes that aim at preventing human injury. The derivation of the concept is described in the following; see Fig. 2. We assume that the human has a certain task-dependent expectation towards the robot's behaviour when both are working in close proximity. For example, the human may expect that the robot moves slowly inside the human's workspace. Our goal is to ensure that the human expectation is fulfilled, which then leads to controlled, intended human behavior instead of possibly hazardous startle and surprise reactions.

The first step in the derivation of the EMU concept is to understand under which circumstances IM occur in HRI. For this, we conduct an experiment where the human reaction is analysed in a common HRI scenario, where the robot approaches the human workspace with variable motion parameters, e.g., speed, acceleration, or direction. The human reaction is recorded and classified via social cue analysis. From the experiments, we derive the relative frequency of IMO depending on the robot velocity and distance between human and robot, which are two relevant parameters that influence IMO (cf. IV-A). We call this mapping a risk matrix for IMO; see 2 c). For a certain scenario and human condition, we can then define a threshold in terms of IMO probability that shall not be exceeded. In the risk matrix, this threshold can be represented by a so-called expectation curve, which relates the current human-robot distance to an expected velocity; see Fig. 2 d). The expectation curve is integrated into the robot motion generation as EMU, which limits the robot speed to a value which is considered expectable by the human if necessary. Finally, the EMU is combined with the Safe Motion Unit (SMU), that provides a safe velocity, which is based on injury data from biomechanics collision experiments. The combination of the two control laws improves safety and trustworthiness of an autonomous system by ensuring that both the human expectation towards its motion is fulfilled and injury is avoided.

In the following, the experimental derivation of a risk matrix and expectation curve is described. Sec. V then considers the implementation and validation of the EMU concept.

IV. EXPERIMENTAL OBSERVATION OF IMO

This section first addresses the prerequisites required for the demonstration of the EMU concept in this paper. Then, the experimental design and procedure as well as the evaluation method for IM are explained.

A. *Prerequisites*

As human expectation is a multi-dimensional problem shaped by various human and environmental factors, a large set of variables influencing IMO exists, which may be considered. For example, the human expectation towards a robot motion depends on the person's situational awareness about an upcoming robot motion, the mental occupancy of the human, and the person's attitude towards technology. Also the robot motion parameters, which may influence the human IMO are numerous, e.g., acceleration, direction and distance of the approach, jerk, and robot velocity.

For the exploratory design of a risk matrix for IMO in this paper, we make the following assumptions. Firstly, to observe the dynamics of IMO in HRI, we can make use of the correlation between expectation and perception shown in [11]. This correlation suggests that the parameters influencing perceived safety affect the IMO as well. Inter-subject distance is a well known factor to human safety perception [24], [26] as well as robot speed [22], [28]. Therefore, the the *instantaneous robot-human distance* d_h and the *absolute robot speed* v_r are chosen in this paper to establish a risk matrix for IMO. Additionally, we consider frontal approaching motion to the human arms and chest. Secondly, we consider the example of collaborative assembly tasks, where the human focuses on her/his task and is aware that the robot approaches to varying distances, but does not await the robot approach at that moment for a specific purpose. The human task may require variable mental occupancy that is measured by *fixation time*. Lastly, to obtain a general understanding of IMO and to validate our approach, we aim for a widely spread spectrum of participants resembling the human factors age, occupancy, and technology affinity.

TABLE II
EXPLORATORY EXPERIMENT DESIGN TOWARDS A RISK MATRIX FOR IMO

Objective	Generate data for expectation curves using maximum reachable velocities	
Approach distance to tablet's edge (see Fig. 3 a):	$\Delta d_{\mathrm{h}1} = 0.00$ m to $\Delta d_{\mathrm{h}6} = 0.25$ m Steps: $\Delta d_{\mathrm{h}} = 0.05$ m Order: randomized	
Approach velocity	Set 1: 0.25 m/s	Set 2: 0.55 - 1 m/s Steps: 0.05 m/s Order: lower v_r with higher d_h
Observed parameter	S-S cues	
Participants	number: 29 age: 34.3 ($\pm$15.9), male: 21 (72.4%), female: 8 (27.6%)	

Following, we conduct an exploratory experiment[1] which models the defined human and environmental condition. By applying a social signal coding scheme, we observe IMO for different v_r and d_h. Resulting, we generate a risk matrix for IMO and expectation curve and embed it to the robot motion generation.

B. *Experimental procedure and design*

The experimental setup introduced in [32] is used. The set-up is depicted in Fig. 3, consisting of a robot manipulator that is mounted on a table, a PC, a camera, which captures the human upper body and face, a tablet placed on a mounting at $0.44\,\mathrm{m}$ distance from the robot base in y-direction and a standing-chair, which ensures that all participants' heads are positioned at approximately the same height in relation to the robot start configuration; see Fig. 3a).

[1]The following experiment was conducted under approval of review number 395/19 S of the Ethics Commission of Technical University of Munich (TUM).

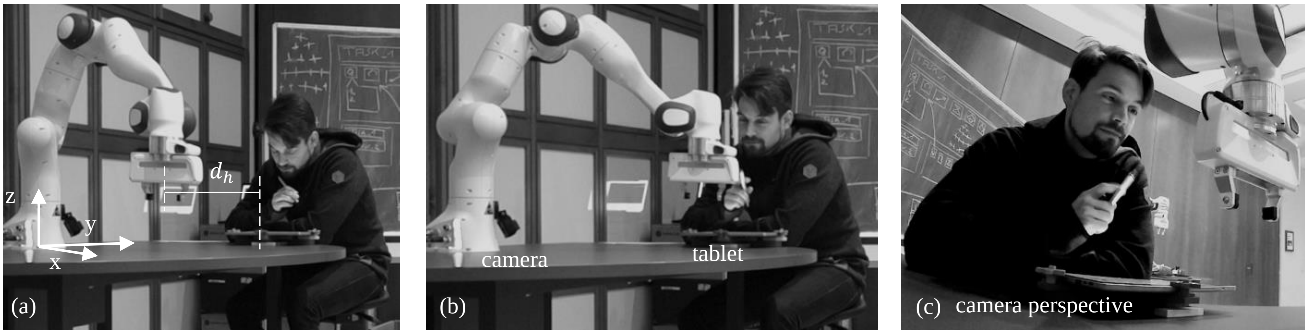

Fig. 3. Experimental procedure: (a) The participant focuses on a quiz while the robot is moving at some distance. (b) The robot approaches towards distance d_h. (c) Mimics and gestures are captured by the camera.

To consider scenarios with few expected IMO a first set of the experiment is defined using $0.25\,\mathrm{m/s}$ as safe velocity applied as general safety limit for teach-in functions of collaborative robots. For a second set the maximum velocities achieved for the corresponding travelling distances are chosen. Table II summarizes the study design. The human-robot distance is defined such that IM may occur, starting with the maximum distance $d_{\mathrm{h6}} = 0.25\,\mathrm{m}$ which is still perceived as comfortable from more than $25\,\%$ of participants in [25] and decreasing to $d_{\mathrm{h1}} = 0\,\mathrm{m}$ in steps of $\Delta d_{\mathrm{h}} = 0.05\,\mathrm{m}$.

The experiment consists of three parts. In the first part, the participant is seated in front of the tablet, and asked to complete a quiz about safe handling of the robot arm on the tablet to obtain mental occupation similar to assembly tasks. The quiz consists of

1. a short fixation time to analyze a problem (focused reading measured among three volunteers: 8.21±4.38 s) followed by
2. a consideration time (e.g. for observation of the environment) and
3. a hand-eye coordination task (tapping on the correct answers).

The participants are told that the robot is moving while they are solving the quiz. The robot is started and the manipulator moves within a randomized number of squares in the x/y-plane at a distance of $d_{\mathrm{h}} = 0.44\,\mathrm{m}$ to the tablet. Then, the end-effector approaches the human workspace and stops at one of six distances ranging from $d_{\mathrm{h}} = 0.25 - 0.00\,\mathrm{m}$ in front of the tablet in a randomized order; see Fig. 3 b). The participant's reaction is recorded by the camera mounted at the robot base. Fig. 3 c) shows the perspective of this camera. During the approach the robot performs a linear movement at an with $v = 0.25\,\mathrm{m/s}$ in the first set and $v = 0.55 - 1\,\mathrm{m/s}$ in the second set as described in Table II. During the approach the robot motion is slightly audible. The participant is not expecting the first approach. For the following five approaches, the participant is aware of the robot motion. Part two serves as training and habituation to the robot and consists of hands-on training in groups of three persons. It starts with an introduction to the *Franka Emika Desk* programming interface and the robot followed by a hands-on programming of a pick and place task. After a short break, the group is split up again and the participants are individually asked to sit down in front of the tablet to complete another quiz. The previous process of robot approach is repeated for every participant.

C. Social-signal analysis for involuntary motion

The experiment requires a reliable evaluation whether the human movement can be classified as IM. Based on the assumption that IM can be measured by social cues of startle and surprise (S-S) as suggested in [32], we conduct a multi-modal video analysis (facial displays, gaze, gestures/postures), in which S-S reactions are recorded and labelled. Even though computational approaches for the automatic recognition of social signals keep on progressing, manual annotation procedures currently are better suited to identify and interpret situated expressions in practical collaboration tasks at first hand [33]. In this article, we use a context-sensitive approach where the expert person evaluates the social signals (human coder) as an interpretation filter, who decides whether a given cue is defined as a social signal of S-S based on theoretical assumptions and sample training. Both inductive sample training and deductive theoretical assumptions informed by already existing knowledge about S-S cues form the basis of our codebook. To achieve a high degree of validity with the maximum possible generalisability of the video analysis, we let two human coders evaluate the video files independently based on a codebook consisting of deductive and inductive codes. The first and main coder is male, has sociological background, and experience in social cue coding. To ensure that his bias is as small as possible, the coder is brought in later to the study, knowing only the general set-up but not the purpose of the study. He is asked to prepare a codebook and to evaluate the videos concerning the frequency and strength of S-S expressions. The second coder is female, mechanical engineer, and has no previous experience in social cue coding. She knows the goal of the experiment and is involved in the programming and set-up. After the first coding the main coder is informed about the study design and the hypotheses. To ensure high reliability concerning individuality and temporal consistency of the annotation, inter- and intra-coder reliability checks are carried out as follows:

- The inter-coder reliability is ensured by using two coders with different levels of experience. Both coders analyse all videos of the experiments and apply the same coding

TABLE III

CODING SCHEME AND RESPECTIVE REFERENCES FOR SOCIAL CUES ON STARTLE AND SURPRISE

	Facial display	Gestures/Postures
Startle (reflex)	- rapid eyeblinks [RE] [35] - lowered eyebrows [LE] [39] - closed eyes [CE] [39] - tightened eyelids [TE] [39] - horizontally stretched lips [HSL] [39] - tightened neck [TN] [39] - delayed felt smile (relief) [DFS]	- evasive head movements [EHM] [36] - evasive trunk movements [ETM] [21] - shoulder jerks [SJ] - body twitches [BT] - body freezes [BF] [21]
Surprise (emotion)	- raised eyebrows [REB] [39] - widened eyes [WE] [39] - raised upper eyelids [RUE] [39] - open jaws [OJ] plus relaxed lips [RL] [39]	- evasive head movements [EHM] [36] - evasive trunk movements [ETM] [21] - body freezes [BF] [21]

instrument. Then, their level of agreement is determined independently.
- The intra-coder reliability is ensured by repetition of the analysis two months later by the main coder and generating the Cohens Kappa score among the annotations [34].

The interpretation of the manual cue coding indicating S-S is guided by the scheme listed in Tab. III. Cues were derived from existing literature [21] and [35]-[36] and an additional expression was identified during sampling training (two videos with a duration of 13 min 53 s in total), which were observed in humans the following unexpected robot movements: felt smiles [20] that appeared to be relief reactions following the immediate S-S responses. For the non-verbal behaviour annotation, we use a simplified version of the MUMIN multimodal coding scheme proposed in [37] and adapt it by including gaze and body postures from human users and removing parts referring to human-human interaction. The ELAN annotation tool [38] is used to classify the video files. As a first step, the coder watches the entire video clip of one participant running through all experimental conditions. Then, the baseline postures, gestures, and facial displays including gaze are annotated at the beginning of the sequence. Subsequently, only the following changes in the expressive behaviour of human participants are annotated, that

- can be associated with movements of the robot arm by registering basic feedback of contact perception ($CP = 1$) and
- which indicate that the human is startled and/or surprised by these movements.

Possible cues for $CP = 1$ include gaze changes towards the robot, which might also be delayed, and interruptions of the human's task, indicated e.g. by freezing hand gestures. If no S-S cue appeared, the IMO was rated as not present.

V. SAFE HUMAN-AWARE MOTION GENERATION

In this section, we present results on the coder reliability for the social signal analysis, the IMO for different robot approach motion trials, and identification of the first risk matrix for IMO. The risk matrix is then used to implement the EMU, which is cascaded with the SMU that ensures physical safety.

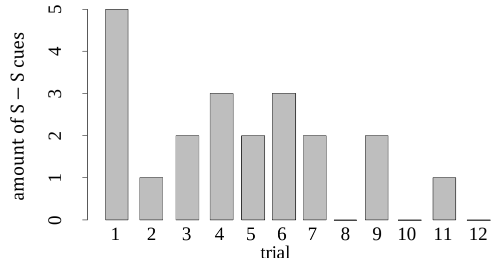

Fig. 4. Total amount of S-S cues with all participants over the number of an robot approaching motion.

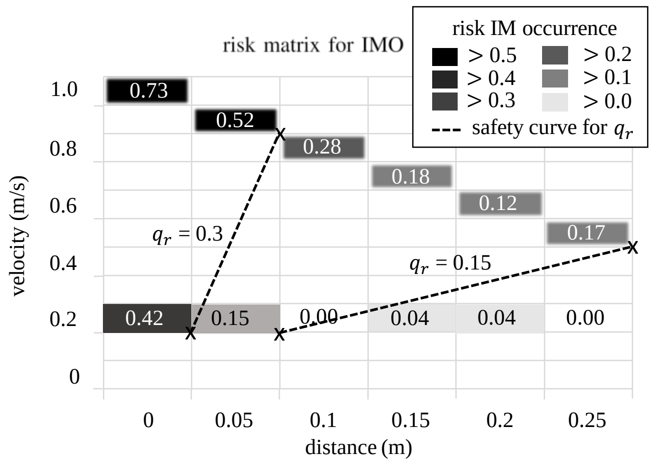

Fig. 5. Exploratory velocity-distance risk matrix for IMO with two exemplary expectation curves for the thresholds of IM $q_r = 0.15$ and $q_r = 0.30$.

A. Reliability of the social signal coding

In order to check the reliability of the social signal coding regarding individuality and time-consistency, we calculate the inter- and intra-coder reliability as explained in Sec. IV-C. The reliability of the evaluation in terms of the Cohens-Kappa score (κ) [34] is as follows:

A) inter-coder reliability: $\kappa = 0.805$
B) intra-coder reliability: $\kappa = 0.840$

This can be considered as "almost perfect" according to [34].

B. Risk Matrix Identification

We obtain the risk matrix for IMO by calculating the relative frequency of IMO within the experiment. For this calculation, we exclude the first trial where the robot approaches the participants, based on the exploratory result of the number of S-S cues identified in the different trials which is shown in Fig. 4. From the significantly higher number of S-S cues on the first trial, we conclude that the participant does not expect the first robot approach motion at all, but is aware of the subsequent approaches.

The experimental results are assembled in a risk matrix, which maps the human-robot distance and velocity to a

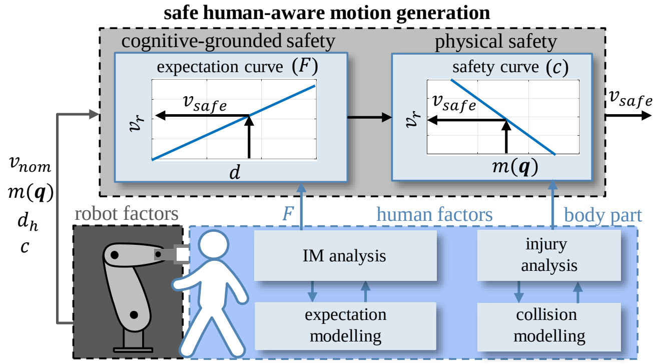

Fig. 6. Safe human-aware motion generation scheme comprising both the SMU and EMU. The framework's inputs are: nominal robot velocity v_{nom}, effective mass depending on the current robot pose $m(\boldsymbol{q})$, instantaneous human-robot distance d_h, currently endangered body part, end-effector curvature c, and human cognitive condition F.

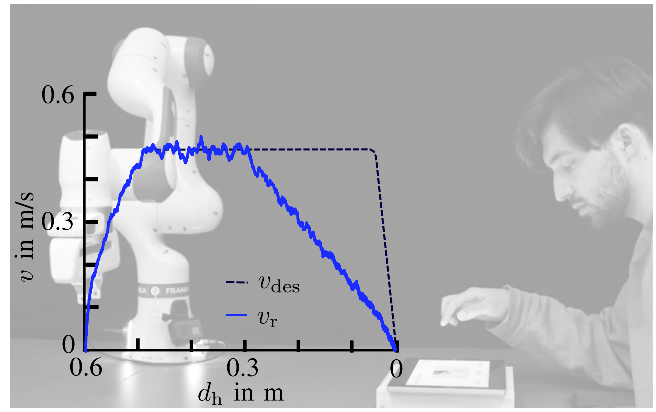

Fig. 7. Measured velocity of the robot following the afe human-aware velocity scaling for the experimental validation compared to the velocity profile using the maximum define task velocity if no human was in the workspace.

probability of IMO; see Fig. 5. In the illustrated risk matrix, the probability of IMO is listed for a) constant velocity ($0.25\,\mathrm{m/s}$) and variable distance (horizontal entries) and b), variable distance and variable robot velocity (diagonal entries). It is notable that the IMO probability decreases at $d_{\mathrm{h}} = 0.10\,\mathrm{m}$ which most likely results from the low number of participants. From the risk matrix, we can deduce an expectation curve for a certain threshold q_r in terms of IMO probability. With this expectation curve, we can determine the maximum robot velocity that can be commanded while satisfying the IMO constraint. Two exemplary expectation curves[2] for $q_r = 0.3$ and 0.15 are illustrated in Fig. 5.

C. EMU Implementation and Validation

In this section, we describe the implementation of the Expectable Motion Unit (EMU) and the validation of the concept. The EMU is cascaded with the Safe Motion Unit [13]. The motion generator is implemented using the Franka Emika Control Interface and the manufacturer's joint velocity controller. For the practical realization of the EMU, we select a linear safety curve for the IMO threshold $q_r = 15\,\%$. Using this safety curve, the instantaneous human-robot distance d_{h} is mapped to the safe velocity v_{EMU}; see Fig. 5. This is done for robot-human distances of less than $30\,\mathrm{cm}$. The velocity limit provided by the EMU is then forwarded to the SMU, which checks whether the desired speed also satisfies the physical safety constraint. The SMU maps the current configuration-dependent robot reflected mas $m(\boldsymbol{q})$ to a biomechanically safe velocity v_{SMU} via a curvature-related human injury threshold called safety curve; see Fig. 7. For the instantaneous robot-human distance $d_{\mathrm{h}} \leq d_{\max}$, the commaned robot velocity is the smallest of the three speeds v_d, v_{EMU}, and v_{EMU}. For distances $> d_{\max}$, the desired velocity is only limited by the SMU.

$$v_{\text{safe}} = \begin{cases} \min\{v_d, v_{\mathrm{SMU}}, v_{\mathrm{EMU}}\}, & d_{\mathrm{h}} \leq d_{\max} \\ \min\{v_d, v_{\mathrm{SMU}}\}, & d_{\mathrm{h}} > d_{\max} \end{cases} . \tag{1}$$

[2]In this paper, the expectation curves are linear for the sake of simplicity. It is possible that they have other shapes.

TABLE IV
RELATIVE FREQUENCIES OF S-S CUES BASED ON ROBOT DISTANCE AND VELOCITY EVALUATED BY EACH CODER WITHIN THE VALIDATION STUDY

d_{h} [m]	v_{r} [m/s]	f by C1	f by C2
0.25	0.405	0.05	0.00
0.20	0.33	0.05	0.00
0.15	0.255	0.00	0.05
0.10	0.18	0.13	0.13
0.05	0.105	0.19	0.14
0.00	0.03	0.25	0.5

To validate whether the EMU concept reduces IMO in practice, we repeat the experimental procedure proposed in Sec. IV-B using the EMU velocity shaping. The experimental setup and and velocity profiles are depicted in Fig. 7. The desired, nominal velocity profile is shown in blue, the velocity that was shaped by the combination of EMU and SMU in green. A group of eleven participants is part of the validation experiment with an average age of 28 ± 4.4 years including six males ($54,5$ %) and five females ($45,5$ %). We expect the relative frequency of IMO observed in this experiment to be less or equal to the desired threshold of $q_r = 0.15$.

The inter-coder reliability for the validation experiment is $\kappa = 0.65$, which can be considered as "substantial" according to [34]. The low inter-coder reliability compared to the previous experiment may be a result of less strong S-S cues (which are difficult to identify) and the lower number of participants. Due to the lower inter-coder reliability, both coders' results are used to verify the EMU. Thus, Table IV lists the relative frequency f of S-S cues rated by both the first (C1) and second coder (C2). According to observations of both coders, we conclude that in case of an approach to a distance $d_h > 10\,\mathrm{cm}$ the relative frequency of S-S occurrence is very low ($0-5$ %). For robot-human distance up to $d_h < 10\,\mathrm{cm}$ the risk of IMO is 13 %. At $5\,\mathrm{cm}$, $14-19$ % IMO is observed. The difference of 5 % of the relative frequency between both coders indicates that the reaction was rated differently in one case. In our preliminary validation we observe that the desired 15 % threshold was satisfied for robot-human distances $> 5\,\mathrm{cm}$. However, for $d_h \geq 0\,\mathrm{cm}$ we observe $25-50$ % IMO.

To sum up, the EMU generated motions that resulted in the desired reduction in IMO in five of six robot approaches. For human-robot distances $\geq 10\,\mathrm{cm}$, the robot velocity can even be increased while ensuring the IMO constraint. However, when approaching the most proximal point of the human body, which was on the boundary of the human workspace ($\sim 4\,\mathrm{cm}$ from the human fingertips), the IMO was not reduced sufficiently.

VI. DISCUSSION

In this paper, we demonstrated the overall approach in an exemplary scenario. All experiments were conducted with a Franka Emika Panda robot arm mounted stationary to a table and under the previously introduced prerequisites in Sec. IV-A. Thus, the results and can not be generalized. Other robot types with e.g. varying size or topology, different scenarios, and different human factors require additional exploratory investigation. To complete the database, multiple experiments are required to define human, robot, and environmental variables influencing IMO and calibration of the safety curves. Here, the importance of the parameter choice lays within the contextual application desired for the EMU, as a full model on human expectation is unfeasible and due changing human state not desirable [10]. Once a set of risk matrices for IMO is established, thresholds for IMO are required, which may either be selected globally, e.g., according probability of the human intruding the robot's workspace or vice versa [40], or dynamically depending on the human condition. For example, one may select a rather low threshold q_r when the user is sleepy, and a higher threshold when the human observes the robot task closely. Therefore, to deploy the EMU approach in real application scenarios, we need to

1) identify and monitor the human condition using a human profiler (eye-tracking can be used to determine the human's level of awareness, for example),
2) select the risk matrix based on a scenario and human condition,
3) define the desired IMO probability threshold and the respective expectation curve.

To fulfill human expectation context-dependent and individualised may also require learning algorithms which can be deployed on top of the general models. Once the robot is capable of safe and expected motion also the human comfort, trust, and acceptance towards a robot may be improved. Finally, a combination of the EMU approach and algorithms for physical human safety enables embedded artificial intelligence that is considered safe and trustworthy.

VII. CONCLUSION

In this paper, we proposed and validated the Expectable Motion Unit (EMU) concept, which aims at avoiding possibly hazardous human involuntary motions (IM) in human-robot interaction. We conducted experiments with 29 volunteers in order to systematically analyse the relative frequency of IM occurrence (IMO) depending on the robot speed and human-robot distance in a typical HRI setting. The experimental results are processed towards a risk matrix, from which safety curves were deduced for a particular application. An expectation curve which limits the probability of IMO to $15\,\%$ in the considered use case was embedded into the EMU motion generator, which limits the robot velocity such that the IM threshold is not exceeded. Furthermore, the EMU was combined with the well-established Safe Motion Unit that ensures physical safety during contact. In a validation experiment, the EMU successfully prevented IM in five out of six cases. Overall, by fulfilling the human expectation towards the robot and taking the biomechanical safety limits into account at the same time, our concept improves the safety, the performance, and the trust of robot users in HRI.

ACKNOWLEDGMENTS

The authors would like to thank Jan Schreiber, Melanie Porzenheim, Leon Oskui, and UAS Fulda for supporting the experiments. This work was supported by the European Union's Horizon 2020 research and innovation programme as part of the project ILIAD under grant no. 732737 and partially supported by the Research Project I.AM. through the European Union H2020 program under GA 871899. We gratefully acknowledge the funding of the Lighthouse Initiative Geriatronics by StMWi Bayern (Project X, grant no. 5140951) and LongLeif GaPa gGmbH (Project Y, grant no. 5140953) and the funding of project KoBo34 (Verbundprojektnummer V5ARA202) by the BMBF (grant no. 16SV7985). Please note that S. Haddadin has a potential conflict of interest as shareholder of Franka Emika GmbH.

REFERENCES

[1] S. Haddadin, A. De Luca, and A. Albu-Schäffer, "Robot collisions: A survey on detection, isolation, and identification," *IEEE Trans. on Robotics*, vol. 33, no. 6, pp. 1292–1312, 2017.

[2] A. Pereira, "Guaranteeing safe robot motion," Ph.D. dissertation, Technische Universität München, 2019.

[3] R. Weitschat, J. Ehrensperger, M. Maier, and H. Aschemann, "Safe and efficient human-robot collaboration part i: Estimation of human arm motions," *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 1993–1999, 2018.

[4] R. Rossi, M. P. Polverini, A. M. Zanchettin, and P. Rocco, "A pre-collision control strategy for human-robot interaction based on dissipated energy in potential inelastic impacts," *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 26–31, 2015.

[5] S. Tadele, T. J. de Vries, and S. Stramigioli, "Combining energy and power based safety metrics in controller design for domestic robots," *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 1209–1214, 2014.

[6] J. Mainprice and D. Berenson, "Human-robot collaborative manipulation planning using early prediction of human motion," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 299–306.

[7] S. Liu, H. Roehm, C. Heinzemann, I. Lütkebohle, J. Oehlerking, and M. Althoff, "Provably safe motion of mobile robots in human environments," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1351–1357.

[8] P. A. Lasota and J. A. Shah, "Analyzing the effects of human-aware motion planning on close-proximity human–robot collaboration," *Human factors*, vol. 57, no. 1, pp. 21–33, 2015.

[9] P. Rani, N. Sarkar, C. A. Smith, and L. D. Kirby, "Anxiety detecting robotic system – towards implicit human-robot collaboration," vol. 22, no. 1, p. 85–95, 2004.

[10] R. A. Knepper, C. I. Mavrogiannis, J. Proft, and C. Liang, "Implicit communication in a joint action." New York, NY, USA: Association for Computing Machinery, 2017.

[11] F. P. De Lange, M. Heilbron, and P. Kok, "How do expectations shape perception?" *Trends in Cogn. Siences*, vol. 22, no. 9, pp. 764–779, 2018.

[12] W. L. Martin, P. S. Murray, P. R. Bates, and P. S. Y. Lee, "Fear-potentiated startle: A review from an aviation perspective," *The Int. Journal of Aviation Psychology*, vol. 25, no. 2, pp. 97–107, 2015.

[13] S. Haddadin, S. Haddadin, A. Khoury, T. Rokahr, S. Parusel, R. Burgkart, A. Bicchi, and A. Albu-Schäffer, "On making robots understand safety: Embedding injury knowledge into control," *The Int. Journal of Robotics Research*, vol. 31, no. 13, pp. 1578–1602, 2012.

[14] *ISO/TS 15066:2016 Robots and robotic devices – collaborative robots*, Beuth-Verlag, Berlin, 2016.

[15] K. Oussama, "Inertial properties in robotic manipulation: An object-level framework," *The international journal of robotics research*, vol. 14, no. 1, pp. 19–36, 1995.

[16] N. Sarkar, "Psychophysiological control architecture for human-robot coordination-concepts and initial experiments," *Proc. IEEE Int. Conf. on Robotics and Automation (Cat. No. 02CH37292)*, vol. 4, pp. 3719–3724, 2002.

[17] R. Alami, M. Warnier, J. Guitton, S. Lemaignan, and E. A. Sisbot, "When the robot considers the human," pp. 1–17, 2011.

[18] H.-L. Cao, G. Van de Perre, J. Kennedy, E. Senft, P. Esteban, A. De Beir, R. Simut, T. Belpaeme, D. Lefeber, and B. Vanderborght, "A personalized and platform-independent behavior control system for social robots in therapy: Development and applications," *IEEE Trans. on Cogn. and Devel. Systems*, vol. 11, no. 3, pp. 334–346, 2018.

[19] S. S. Tomkins, "Surprise–startle: The resetting affect." *Affect, Imagery, Consciousness*, vol. 1, pp. 498–522, 1962.

[20] P. Ekman and W. V. Friesen, "Felt, false, and miserable smiles," *Journal of nonverbal behavior*, vol. 6, no. 4, pp. 238–252, 1982.

[21] M. Coulson, "Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence," *Journal of nonverbal behavior*, vol. 28, no. 2, pp. 117–139, 2004.

[22] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International journal of social robotics*, vol. 1, no. 1, pp. 71–81, 2009.

[23] T. J. Wiltshire, E. J. Lobato, D. R. Garcia, S. M. Fiore, F. G. Jentsch, W. H. Huang, and B. Axelrod, "Effects of robotic social cues on interpersonal attributions and assessments of robot interaction behaviors," *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, vol. 59, no. 1, pp. 801–805, 2015.

[24] E. T. Hall, *The hidden dimension*. Garden City, NY: Doubleday, 1966, vol. 609.

[25] M. L. Walters, K. Dautenhahn, R. Te Boekhorst, K. L. Koay, C. Kaouri, S. Woods, C. Nehaniv, D. Lee, and I. Werry, "The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment," *Proc. IEEE Int. Workshop on Robot and Human Interactive Communication (ROMAN).*, pp. 347–352, 2005.

[26] L. Takayama and C. Pantofaru, "Influences on proxemic behaviors in human-robot interaction," *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 5495–5502, 2009.

[27] J. Mumm and B. Mutlu, "Human-robot proxemics: physical and psychological distancing in human-robot interaction," *Proc. of the 6th Int. Conf. on Human-Robot Interaction*, pp. 331–338, 2011.

[28] D. Kulić and E. Croft, "Anxiety detection during human-robot interaction," *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 616–621, 2005.

[29] C. K. Or, V. G. Duffy, and C. C. Cheung, "Perception of safe robot idle time in virtual reality and real industrial environments," *Int. Journal of Industrial Ergonomics*, vol. 39, no. 5, pp. 807–812, 2009.

[30] S. Nonaka, K. Inoue, T. Arai, and Y. Mae, "Evaluation of human sense of security for coexisting robots using virtual reality. 1st report: evaluation of pick and place motion of humanoid robots," *Proc. IEEE Int. Conf. on Robotics and Automation(ICRA)*, vol. 3, pp. 2770–2775, 2004.

[31] F. Dehais, E. A. Sisbot, R. Alami, and M. Causse, "Physiological and subjective evaluation of a human–robot object hand-over task," *Applied ergonomics*, vol. 42, no. 6, pp. 785–791, 2011.

[32] R.J.Kirschner, L. Burr, M. Porzenheim, H. Mayer, S. Abdolshah, and S. Haddadin, "Involuntary motion in human-robot interaction: Effect of interactive user training on the occurrence of human startle-surprise motion," in *2021 IEEE/ISR International Conference on Intelligence and Safety in Robotics (ISR)*, 2021.

[33] K. Pitsch, "Limits and opportunities for mathematizing communicational conduct for social robotics in the real world? toward enabling a robot to make use of the human's competences," *AI & SOCIETY*, vol. 31, no. 4, pp. 587–593, 2016.

[34] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *biometrics*, pp. 159–174, 1977.

[35] C. Grillon and J. Baas, "A review of the modulation of the startle reflex by affective states and its application in psychiatry," *Clinical neurophysiology*, vol. 114, no. 9, pp. 1557–1579, 2003.

[36] M. Davis, "The mammalian startle response," *Neural mechanisms of startle behavior*, pp. 287–351, 1984.

[37] J. Allwood, L. Cerrato, K. Jokinen, C. Navarretta, and P. Paggio, "The mumin coding scheme for the annotation of feedback, turn management and sequencing phenomena," *Language Resources and Evaluation*, vol. 41, no. 3-4, pp. 273–287, 2007.

[38] P. Wittenburg, H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes, "Elan: a professional framework for multimodality research," *Proc. 5th Int. Conf. on Language Resources and Evaluation (LREC)*, pp. 1556–1559, 2006.

[39] P. Ekman, W. V. Friesen, and R. C. Simons, "Is the startle reaction an emotion?" *Journal of personality and social psychology*, vol. 49, no. 5, pp. 1416–1426, 1985.

[40] E. Kim, R. Kirschner, Y. Yamada, and S. Okamoto, "Estimating probability of human hand intrusion for speed and separation monitoring using interference theory," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101819, 2020.

