Text accepted (02.05.14) for publication in Proceedings of the 23th IEEE Ro-Man 2014. [PREPRINT]

A robot as fitness companion: towards an interactive action-based motivation model*

Luise Süssenbach[1,2], Nina Riether[1], Sebastian Schneider[1], Ingmar Berger[1]
Franz Kummert[1], Ingo Lütkebohle[1], Karola Pitsch[2]

***Abstract*— The topic of *motivation* is a crucial issue for various human-robot interaction (HRI) scenarios. Interactional aspects of motivation can be studied in human-human interaction (HHI) and build the basis for modeling a robot's interactional conduct. Using an ethnographic approach we explored the factors relevant in the formation of motivation-relevant processes in an indoor–cycling activity. We propose an interactive, action–based motivation model for HRI that has been implemented in an autonomous robot system and tested during a long-term HRI study. The model is based on micro-analyses of human indoor cycling courses and resulted in an adaption of specific dialog patterns for HRI. A qualitative evaluation – accompanied by a quantitative analysis – demonstrated that the transfer of interaction patterns from HHI to HRI was successful with participants benefitting from the interaction experience (e.g., performance, subjective feeling of being motivated).**

I. Introduction

Doing sports is crucial for health and well–being. However, a range of factors in our daily lives often prevent us from exercising regularly and thus, point to the importance of *motivation*. In particular, during a workout people tend to be more motivated and perform better if they are accompanied by a personal trainer [20]. This is all the more relevant for people working in closed environments (e.g., astronauts on long-term space missions). They are faced with multiple physiological and psychological challenges, which is why daily workout is particularly important. However, in closed environments such as a space stations motivating *real–time* feedback during exercising is often impossible because of technical constraints (e.g., time delays for ground communication, etc.). Hence, robotic or virtual support might be useful, because it can be situated and tailored to the user's performance.

Nowadays, sociable robots/agents are being explored for supporting humans in quite diverse daily situations (e.g., sport training, robot assisted weight loss, tutoring) [13], [4], [8], [2], [3]. In general, motivation is studied by psychologists with a focus on the individual level. However, motivation in a two (or more) party situation is also an interactional phenomenon that involves multimodality and thus, has interactional aspects that can be especially well studied in HHI [12]. Using an ethnographic approach we explored the factors relevant in the formation of motivation–relevant processes in an indoor–cycling activity. The qualitative analysis of this HHI resulted in an interactive, action-based motivation model for HRI that aims at emulating the communicative resources and sequential flow used by a human instructor. This model was then implemented on the NAO robot that acted as an indoor cycling instructor. A long-term HRI study (18 consecutive days in which participants where isolated) was conducted, in which Nao acted as a sports companion who individually accompanied 8 participants during their daily 50 minutes indoor cycling workout session.

The current paper addresses the following research questions: (i) How can motivation–relevant processes be captured and illustrated in an interactive, action–based model?, (ii) How does an interactive action–based motivation model developed from HHI perform in HRI? and (iii) How does a robot system equipped with this model perform in comparison to a non–reactive system?

In what follows we will present qualitative findings of HHI that lead into a model for motivation on an interactional level. This model in an HRI situation will be presented in depth with a single-case analysis. This qualitative analysis is followed and embedded in quantitative results.

II. Background: Motivation in Interaction

Over the past decades, research in psychology has contributed to our understanding of motivation (e.g. [11], [21]). Motivation is understood as a force that drives human behavior and which should therefore also play a crucial role in a range of HRI scenarios.

A. Motivation in studies on human interaction

From a sociological and linguistic perspective, motivation is not only considered as the *internal* state of *one* individual, but also as a collaborative achievement of two or more persons interacting with each other. Thus, motivation constitutes also an interactional process carried out in real–time and mobilising the multimodal communicational resources of the participants. A recent field study [12] pointed out that physical coordination and corporeality (e.g. boxing) are crucial for keeping motivation in sports. Moreover, an ethnographic video-based study in gyms highlighted that instructors and trainees make use of different repair strategies to coordinate each other and to benefit from the workout [25], [23]. This finding indicates that responsiveness and reactivity – realized through repair mechanisms – play an important role for

*We gratefully acknowledge support by the German Aerospace Center (DLR) - support code 50RA1023 - with funds from the Federal Ministry of Economics and Technology (BMWi) due to resolution of the German Bundestag. Karola Pitsch acknowledges the financial support from the Volkswagen foundation (Dilthey fellowship).

[1]L. Süssenbach (lsuessen@techfak.uni-bielefeld.de), N. Riether, I. Berger, F. Kummert, I. Lütkebohle is with the Applied Infomatics Group.

[2]K. Pitsch and L. Süssenbach are with the Interactional Linguistics & Human-Robot-Interaction Group.

understanding motivational processes. Also in everyday life people are faced with the task of expressing themselves in a way that establishes the desired communication with the address. The competences necessary for such a task include knowledge about how to avoid misconceptions in advance as well as knowledge about how to deal with communication that turned wrong. Hence, repair strategies like self-initiated as well as other-initiated repair play an important role for a reciprocal, goal-oriented achievement of interaction [19].

B. Motivation and socially assistive robots

Motivational aspects in interaction are complex processes that are on the one hand relevant in a single short–term interaction, but also particular crucial in long–term repeated interactions. A longitudinal control group study [2] in which participants interacted with a fitness-tracking virtual rational and non-rational agent showed that a rational designed agent got a higher human-like ascription (more respect, kinder, etc.) Going beyond virtual agents, past research demonstrated that an embodied robot system is able to support humans in cognitive [18], [9] and assistive [8] tasks. Furthermore, Powers et al [17] found that a robot had more social impact than a computer agent. In this study, although whether participants interacted with a robot or with an agent had no influence on health behavior, participants spent more time with the robots compared with the agents. Powers and colleagues demonstrated that robots can be efficient communicators that are perceived as helpful and useful. Indeed participants may even attribute positive personality trait ratings and lifelikeness to them. Furthermore, Fasola et al. [3] conducted a HRI motivation study with a control group design, investigating praise and personalized address. The robot demonstrated a body movement and then instructed the participant to imitate it. The robot gave real–time feedback, providing corrections during imitation as well as praise. Here, participants interacted only once for 20 minutes with the robot, therefore long-term effects in robot assistive sport interaction could not be investigated. Only few studies have investigated long-term interaction effects so far (e.g. [9], [7], [25]). Most of these HRI long-term studies revealed that the robot failed to keep the participant's interest after the first interactions, mostly because the first impact created unreasonably high expectations. Slow response times, incorrect reactions as well as low interactivity seem to be responsible.

In this paper, we will focus on the ways in which interactional processes between user and robot affect the users' inferences concerning the robot's capabilities and competences. To do so, the present study used repeated and longer periods of interaction in which the user perceived the robot more intensively and had more time for interacting with it.

III. INTERACTIONAL STRUCTURES IN HHI

A. Motivation in a real–world Instructor–Trainee Interaction

To develop a robot system that adequately motivates participants, we started with HHI observations of workouts guided by four indoor cycling instructors. Information concerning the macro-structural level as well as on the level of detailed micro sequential structures were gathered. Thus, in a first step we investigated how an exercise unit in a real-world indoor cycling course is set-up. Sessions generally began with a greeting, followed by the physical exercise and finally ended with small talk and a farewell. Within these phases sub-structures that sports literature refer to as *movements* can be identified. Movements describe intervals in which different types of exercises are practiced[1]. To investigate motivational relevant processes between instructor and trainee more systematically , we conducted a semi experimental study, in which specific conditions in their natural environment has been manipulated (e.g., instructor has no bike, no music)[2].

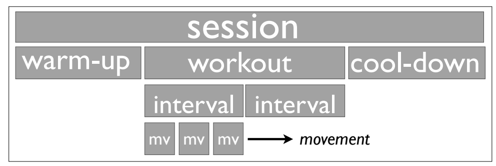


Fig. 1. Typical structure of an Indoor–Cycling workout

Our analysis showed that the interaction between instructor and trainee is strongly sequential. First of all the instructor initiates the movement via a *preparation*, a short verbal statement that attracts trainees' attention and marks a moment in time, at which the exercise will change in a relevant way. So, semantically, these preparations often use time markers (e.g., "shortly", "in a few seconds") and short statements that illustrate how the exercise will change (e.g., "it goes uphill."). The preparation is not obligatory. After the preparation an *instruction* is given. The instruction is often a short verbal statement that mainly functions as a time marker (e.g., "hep!", "hop!"), but it can also include practical advice (e.g., "up!", "down!"). In the following fragment the instruction is much more syntactically and semantically extended. The instructor marked the upcoming exercise as difficult (line 01).

```
[...]
01  I-ver:  so, approach this carefully
            also taste dich mal so ran

            so that you get this slow rhythm
            dass du auf diesen la::ngsamen rhythmus runterkommst
                                                     Instruction
```

MUSIC STARTS

Up to this point the interaction sequence is mostly uni–directional. The instructor permanently monitors the trainee, and acts only according to the workout schedule. If the trainee follows the current instructions the instructor uses the *continuer feedback* to keep the trainee motivated. This feedback is a simple statement which signals that the instructor recognizes the trainee's effort and compliance and serves to strengthen the trainee's motivation. If the trainee does not follow the current instruction the instructor reacts using a *repair strategy*. In the following fragment the instructor intervened by saying "increase your resistance" (l.03).

[1] such as training different muscle groups vs. training the cardio-vascular system

[2] for details please refer to [23], [24]

```
02  I-gaz:  .....@flywheel.......
            Trainee does not follow
03  I-ver:  increase your resistance
            und- dreh noch weiter
                                  Repair
```

Repair sequences consist of either gaze organization, gesture or a verbal repetition of the current instruction with a changing intonation ("UP!") or a short statement ("pedal faster!"). As soon as the trainee is compliant the instructor reacts with a *positive feedback* (e.g., "perfect!", "great job!") – short statements that strengthen the trainees' current action.

```
04  I-gaz:  |....................@flywheel........................|
    I-ver:  there's the problem to overcome the resistance, isn't it?
            da is schon das problem über die mitte rüberzukommen; ne?
                          Trainee still does not follow = Failure
05  T-ver:  ah
            ah::
06  I-ver:  but because of the flywheel you can see perfectly
            aber man sieht su::per an der scheibe,
              *Img. 22:44.530
            that you've got power
            dass du richtig power drauf hast
            [...]                          Encouragement Feedback
07  I-ver:  very positive sign
            seh::r gutes zeichen.
            Encouragement Feedback
```

#Img 22:44.530

If the trainee is still not able to follow the instruction the instructor makes use of *encouragement feedback* - statements that are rather different from other means employed by the instructor because of length and syntax. Taken together, it can be noted that repair strategies are an important topic in instructor-trainee interactions: They keep the exercise routines fresh and interactive, prevent the trainee from exercising incorrectly and keep the trainee motivated during high physical load phases.

To sum up, in human–guided indoor cycling courses the instructor permanently monitors the trainee's performance (e.g., cadence, rhythm to music, physical load via face color etc.), instructs him/ her using short verbal statements and evaluates the execution. If a repair is necessary the instructor uses several repair strategies ranging from a repetition to a changing gaze orientation. Moreover, it is noticeable that the instructor's verbal statements are mostly short and elliptical. Moreover findings also indicate that the facial expression of the instructor plays a crucial role. If more power is required by the trainee the instructor often smiles. The workout itself is not a *uni–directional*, but rather a reciprocal multimodal *inter*-action between instructor and trainee. These findings in human–guided workouts suggest that a robot system in the role of a fitness instructor needs a fine-grained and sequential model for a motivating interaction experience.

B. From HHI to HRI: towards an action-based motivation model

The qualitative analysis revealed a complex multimodal structure of motivation–relevant processes that are fine-grained and sequentially organized within the instructor–trainee interaction (e.g. gaze organization, prosody, gesture). Because a robot system is limited in its capabilities a reduced model is required. This action–based motivation model (s. Fig.2) represents the sequential structure of a movement and represents all sequential possibilities within an interaction. As a premise, instructor (I) and trainee (T) interact face-to-face, thus enabling mutual monitoring activities.

The movement starts with a preparation utterance and is followed by the instruction. At this point, either the user *is able* to realize the instructed exercise or *is not able* to do so. If the user is able, the robot verbalizes the continuer feedback. If the user is not able to put the instruction into practice, the system provides corrective feedback via repair advices (e.g., pedal faster, slow down, get up, etc.). The sequence ends with an evaluation utterance that leads either to a positive or an encouragement feedback. Whereas positive feedback is a very short statement, an encouragement feedback utterance is syntactically and semantically extended.

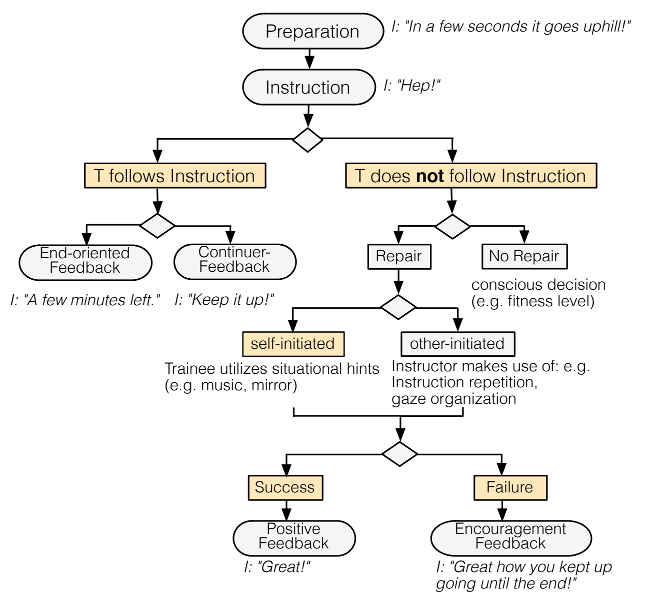


Fig. 2. Interactive action-based motivation model: Instructor (I) in grey; Trainee (T) in orange

C. Communicative Resources in HHI

Besides the sequential verbal structure, the multimodal resources the instructor uses in interaction are of utmost importance. The qualitative analysis of our video–recorded indoor cycling courses have shown that the instructors often use similar verbal realizations of i.e. *unspecific preparation-type* ("attention!"), only changing the intonation, prosody and accentuation. The multimodal resources range from different pointing and body gestures, to gaze orientation to nodding activity [24]. The individual sequential outputs are not verbal

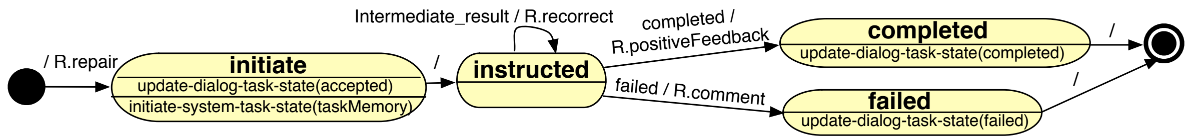


Fig. 3. The Robot Correctable Instruction (RCI) dialog pattern: R.repair = e.g."Pedal faster!", R.recorrect = e.g."Come on!", R.comment = e.g."Try it next time."; The RCI triggers multimodal robot utterances (BML) (e.g. R.repair = *"Pedal faster!"* + *Gesture* + *colored LEDs*).

statement at all, but rather *multimodal*, holistic linguistic characters that include a body of different communicative resources.

IV. Transforming the action-based model to an Interaction Pattern for HRI

To rebuild repair strategies from human–guided workouts we formulized a dialog pattern called the *Robot Correctable Instruction* (RCI)[3] using the PaMini dialog system [14], which is a framework for designing mixed–initiative HRI based on generic interactions patterns. The RCI pattern describes an action initiated by the robot. The positive course of events is that the robot requests an instruction to be executed, the dialog manager initiates the domain task, the responsible system component accepts execution so that the dialog manager will assert execution. Finally, the task is completed and the robot acknowledges this completion: On an interactional level this means, that the robot system recognizes the present deviation regarding the instructed value, triggering the RCI pattern. That initiates the repair advice (e.g. "Pedal faster" + Gesture). The system checks the *intermediate result* and gives positive feedback (e.g., "Great! + Gesture", positive way), again corrective feedback (e.g., *"Come on! + Gesture"*, retry) or failed and ends the task. This pattern enables us to rebuild human repair strategies and to maximize beneficial effects for participants (e.g., avoid false workout execution). Because multiple errors are possible on side of the trainee (posture, resistance, cadence), the robot system follows a repair hierarchy structure that takes the character of an exercise into account – 1. posture, 2. resistance, 3. cadence.

A. Textual Adaption for HRI

For comparability reasons (with the control group) the robot verbalizes within the instruction concrete values (regarding watt, cadence, sitting, standing) that the participant has to achieve. In the control group participants received the same instructions with concrete values textual via a display.

B. Synchronizing the robot's communicative resources

Regarding the multimodal design of each verbal utterance, we used the Behavior Markup Language (BML) [10] that allows a multimodal synchronized generation of several communicative resources (i.e. verbal statements, prosody, gesture, color of LEDs) that enabled us to rebuild human–like fitness advices using different multimodal channels of communication.

[3]for conventions please refer to [14]

C. Overview of verbal utterances

To design the complete workout over 18 sessions (including the interactional framing, preparation and instruction sequences as well as repair situation) we used 3 different dialog pattern (including RCI) with 112 textual different configurations in sum that in total included 50 preparations, 86 instructions, 66 repair utterances as well as 22 welcome and farewell statements.

V. Robot System

The ethnographic analysis of human fitness instructors revealed that interaction during indoor cycling is fine-grained and dependent upon the trainee's actions. We used these observations as a guideline to build an autonomous system that guides users through a workout and gives them appropriate feedback based on their execution over extended periods of training sessions. As the robotic target platform we used the humanoid Nao[4]. The design and implementation of such system comes with a variety of challenges we needed to cope with. The system needed to be perceptive to recognize the user's training execution, vital data and own position, it furthermore needed to be able to make decisions based on these parameters, it needed to be reactive to put these decisions into actions using dialog systems and multimodal feedback (e.g. speech, gesture, head orientation and color changing of eye-LEDs) to give the user corrective or positive feedback. Due to the non-technical focus of this work, we only sketch the implementation and component details of the system. The robot's behavior during scenario-specific workout situations is triggered by the *action-based motivation model* designed as a state chart [5], [1]. The usage of state charts allows us to build reconfigurable patterns that can be adapted to the different exercises we encounter in indoor cycling regimes. States trigger situation–specific interaction patterns, which are designed in our dialog sytem PaMini [14]. These interaction patterns then execute multimodal behaviors modeled as the BML (cf. Sec. IV). Besides the embodied vision for face detection and marker detection for localization, the robots not embodied perception for detecting the user's physical state and condition. These are:

a) Bike Computer: Participants exercised on an indoor bike from SRM[5] which provides via the Powercontrol accurate values for current cadence, power and speed. This enables the system to detect deviating values and to react in an adequate way.

[4]http://www.aldebaran-robotics.com/
[5]http://www.srm.de/

b) Posture and Pedal Detection: To assess the participants' posture on the bike, two 3D depth cameras were used, one in front of the bike and one beside the bike (s. Fig. 3). The posture component was used to identify whether the user was sitting or standing.

c) Heart Rate Monitor: In order for the robot to evaluate the participant's performance and to detect the physical limit the robot needs to know the participant's heart rate using a heart rate belt. This allowed us to record the heart rate variability also.

VI. Data & Study Set-up

During the experiment, 96 workout sessions of HRI were recorded with two net cameras. One camera has been positioned in front of the participant in order to record the partcipant's facial expression and gaze. The other camera has been set up right-side behind the participant to record the whole interaction scene between participant and robot. In addition, we logged performance data of participants (watts, heart rate, cadence). The robot Nao was positioned on a table in front of the participant.

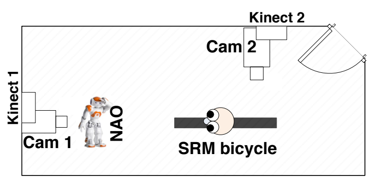


Fig. 4. Experiment set–up: NAO as an indoor cycling instructor

VII. HRI Study Procedure

To investigate how the interactive action–based motivation model performs in HRI we conducted an 18-day randomized, controlled isolation study in which 16 participants were tested in groups of 8 participants each. The aim of the study was to simulate conditions of manned long–term space missions, with one group being accompanied by a robot assistance system and the other by a textual control system (control group). The daily activities and schedules of the two groups were identical. One of those activity was doing sports. Participants were asked to cycle every day, thus 18 times. Participants in both groups completed identical workout schedules, i.e. they underwent equal exercises concerning the physical performance effort necessary. In the control group participants viewed a computer display that visualized all relevant information in written form. If the current exercise changes a beep sounded. The display system was neither interactive nor responsive - thus, it provided instructions and structural information only. Participants that trained with the robot system were welcomed by the robot, took part in an interactive and responsive workout, that, besides giving structural information and instructions, also provided repair hints as well as positive and/ or corrective feedback.

A. Partcipants and Design

16 participants (all men) with an average age of 23.63 years took part in the 18-day isolation study. The control group (text display) and experimental group (robot system) consisted of 8 participants each, who were matched for personality and physiological parameters. After an extensive pretesting phase, potential participants with extreme values on personality characteristics or physical fitness as well as persons with prior experience in robotics were excluded from participation. All participants were healthy and had a Body Mass Index (BMI) between 20-25 kg/m^2. All participants were non-smokers and successfully completed the medical as well as the psychological qualification. They received monetary compensation for their participation. The study was approved by the Ethics Committee of the Deutsche Gesellschaft für Psychologie (DGPs; German Psychological Society).

VIII. Analytical Method

In order to investigate the influence of the robot's behavior on the way in which the participant perceived the robot, we used an analytic approach taking into account the sequential structure of the interaction. This approach is based upon Conversation Analysis (CA). CA is focused on the order and sequential patterns of interaction and the micro-coordination between the actors and detects analytical phenomena [22]. Thus, this approach allows us to investigate the interrelationship between the robot's and trainee's actions and how they respond to each other. CA is especially apt for our research purposes as it allows us to *reconstruct* the participant's view. More specifically it provides us access to the trainee's perception of the robot's actions and thus enables an analysis of *how* and especially *why* the perception changes during interaction. We start with an explorative, qualitative analysis of one single case to detect the organizational features and interactional phenomena in order to reveal the relevant analytical issues and categories using CA. CA's fine-grained analysis of interaction sequences demands a repeated inspection of video-taped data and the transcription of all relevant events. The aim is to highlight the structural organization of actions (e.g., changing gaze, noticeable prosody) and to clarify how one action is contingently relevant for another action. Due to these structural properties, also the absence of an (otherwise expected) action can become relevant. Moreover, in this paper we combine qualitative, CA-based analyses with quantitative analytical methods.

IX. Qualitative Evaluation: Motivation & Repair

To evaluate how the model performed during the HRI study, we need to investigate the interplay of the systems's and the user's actions during the workout. Therefore, we will present a qualitative analysis of two single cases. The first fragment represents a typical repair sequence that can be captured through all HRIs in the corpus (96 workouts). The sequential analysis reveals the fine-grained sequential interrelationship between participant's and robot's action and explore *how* the trainee makes use of the robot's repair strategy and puts the advices into beneficial practice. The fragment[6] is from session 3 of participant P21, when user is already familiar with the system The situation is located at the very beginning of the

```
Participant & Session No.: P21, Session 3
Fragment: 04:03 - 04:30
01 R-ver: we warm up
          wir wärmen uns auf

          with cadence of 60 and 57 watts
          mit SECHzig umdrehungen und siebenundfünfzig watt
                                                  Instruction
02 P21 RESISTANCE > TARGET VALUE
        Trainee does not follow

03 R-ver: your resistance is too high. |resistance minus, minus.
          du hast zu viel widerstand.    widerstand minus, minus.
   P-act:                              |-res.....     Ist Repair
```

workout, the warm–up. The robot has instructed pedaling with a cadence of 60 with 57 watts (l.01). However, observed resistance is too high and cadence is too low. Hence, the *Robot Correctable Instruction* is triggered. Because of the repair hierarchy the robot first repairs resistance, verbalizing "your resistance is too high. resistance minus minus" (l.03). With the repair the trainee changes his view and looks at the robot. Immediately, the trainee re–adjusts the resistance (l.03). Then,

```
04 P21 RESISTANCE > TARGET VALUE
        Trainee does not follow
05  R-ver: pedal faster (1.0)
           tritt schneller
                        2nd Repair
06 P21 RESISTANCE + CADENCE = TARGET VALUE
                        Trainee does follow
07 R-ver: goody.this way it's correct.
          gut. SO is es richtig.
                        Positive Feedback
```

the next repair advice regarding the cadence is verbalized by the robot ("pedal faster!"). The participant reacts promptly and pedals faster. The fragment ends with the positive feedback utterance "goody. this way it's correct." (s. l.07). The analysis of fragment 1 illustrates that the implemented model for motivation and interaction based on repair–sequences seems to be functional. If the participant is not able to follow, the RCI pattern is triggered. Over the entire corpus participants make use of the repair advices. Thus, the workout between robot and participant flows and the participants benefit from the robot-supported workout.

However, in HRI situations people often need time to adapt to the unfamiliar robot system. A number of studies pointed out that in the initial phase of an interaction, humans quickly form still vague ideas concerning the robot's capabilities [16], [25], [15]. This is also true for the present HRI situation. Whereas the fragment presented above was from session 3, after the participant has already interacted with the system for 50 minutes each workout, the upcoming fragment is located at the very beginning of session 1. It illustrates a disturbed interaction situation and points to the difficulties that are likely at the beginning of a HRI. The session has started for about 10 minutes and the robot initiates one of the first exercises. Similarly to the first fragment, the robot starts the movement by verbalizing the preparation statement "attention" (l.01).

[6]for Transcription Conventions s. Appendix

Immediately, the participant looks to the robot for about

```
Participant & Session No.: P21, Session 1
Fragment: 18:38-19:53
01 R-ver: attention(-)
          achtung
          Preparation

02        get out off the saddle (3.0)
          GEH aus dem sattel
                            Instruction
03 R-ver: with 151 watts
          mit einhunderteinfünfzig watt.
                            Instruction

04 PARTICIPANT IS STILL SITTING FOR 5 SEC.
                   Trainee does not follow

05 R-ver: get up -
          steh auf
          Ist Repair

06 R-ver: sit down (15.0)
          setzen
```

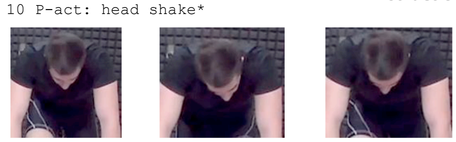
one second, but quickly orients back, looking down. The robot continues and gives the first instruction *"get out off the saddle"* (s. l.02). At the end of the instruction the participant changes his gaze and looks to the robot. He monitors it for about 3 seconds, but does not follow the instruction. Meanwhile the robot gives the next instruction *"with 151 watts"*. Neither does the participant get out off the saddle nor does he try to achieve the requested watts in another way (e.g. increasing cadence). The robot does not react to the

```
04 P21 IS STILL SITTING
         Trainee does not follow
05 R-ver: get up- (1.0)
          steh auf
                Ist Repair
06 R-ver: sit down
          setzen
```

discrepancy, but continues with the next exercise. Now the robot instructs the trainee with *"get up"*, and then does so quickly again *"sit down"*. Here, *"sit down"* is realized after posture recognition. However, the robot system recognizes the participant's posture wrongly. Here, it becomes obvious, that internal information and outer recognitions clash. Until now, on an interactional level the system failure is not apparent and the participant does not treat it as such. The two instructions

```
07 R-ver: get up -
          steh auf
           2nd Repair
08 P-act: ↑  (1.5)

09 P-act: ↓  (.) +res | ↑  (3.0)| |  ↓    | (-)
   R-ver: |continue seated || with 151 watts  (-)|
                                        Instruction
10 P-act: head shake*
```



*19:52.012 *19:52.311 *19:52.673

could be a simple posture change. As before, the participant does not follow the instructed exercise. Within this movement the participant does not realize any requested exercise. 15 seconds pass till the robot produces the next utterance. Again, the utterance "get up" is a posture recognition failure, but functions in the situtation as a repair (l.07). Immediately, the participant gets up, but quickly sits down again (l.08).

The bicycle's resistance seems to be too low, therefore he pedals in vain. Hence, he gets back in order to increase the resistance (l.09). Afterwards he gets up again to follow the instruction (l.09). Simultaneously when he gets out off the saddle the robot instructs *"continue seated with 151 watts"*. Here, user expectation and system conduct mismatch. Despite problems to realize the robot's instruction, the participant has tried to do so. Moreover, when he got the bike setting correctly and intended to continue, the upcoming instruction of the robot finally aborted participant's attempt. Obviously, this mismatch leads to dissatisfaction and incomprehension: promptly, the participant looks down and shakes his head hard (l.11).

Both fragments illustrate that participants need time for reconstructing the sequential flow of the robot-supported workout. With passing time the participant uses the robot as an instrument to help him undertake the workout and is able to put the instruction and repair strategies into beneficial practice. This indicates the strong relevance of studying repeated long-term interaction patterns. However, further qualitative investigations with a bigger sample are needed.

X. Quantitative Evaluation: Physical & Subjective Measures

As quantitative measures, physical variables (development of resting heart rate) as well as subjective questionnaire data (partly adapted from NASA TLX scale ([6]), i.e. perceived physical training requirement, training enjoyment, physical effort during training, training motivation) were recorded throughout the study. A first preliminary analysis of the quantitative data supports the qualitative video–based findings. The analysis revealed that the resting heart rate (measured

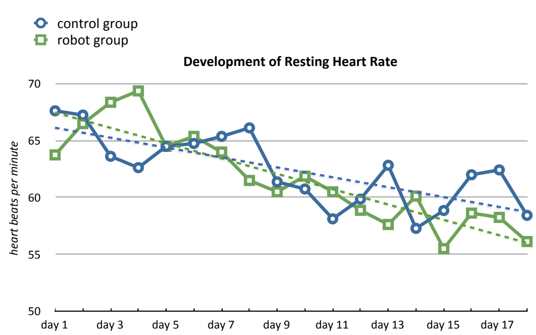


Fig. 5. Development Resting Heart Rate

daily after wake up) significantly decreased over the course of the study both for the experimental group training with the robot ($-0.675x + 68.151$, $p < .001$) as well as the control group ($-0.476x + 66.752$, $p < .01$), which shows that both groups profited from daily training, leading to increased physical fitness (s. Fig.6). This training effect was marginally enhanced in the robot group compared to the control condition ($p = .053$) - in light of the small sample size of 8 participants per group all the more meaningful. Concerning subjective data obtained via questionnaires after every training session, training enjoyment did not differ between the two groups ($p > .05$) and also remained constant over the course of the

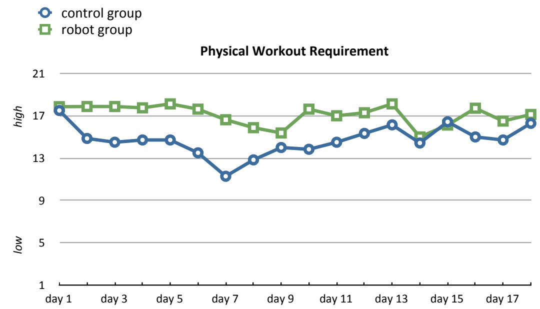


Fig. 6. Physical Workout Requirement

study (robot group: $-0.008x + 4.587$, $p = .723$; control group: $0.002x + 4.187$, $p = 0.939$). However, participants in the

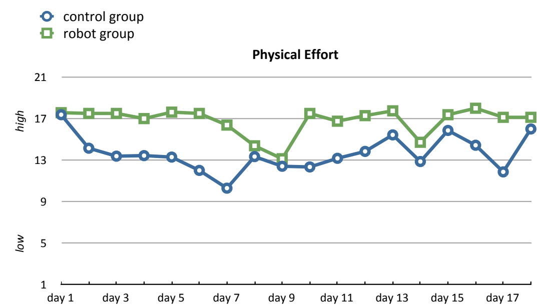


Fig. 7. Physical Effort

robot group perceived the physical training requirements as significantly more challenging compared to the control group ($t(14) = 4.306$, $p < .001$, *Cohen's* $d = 2.302$, one-sided), even though both groups followed the exact same training regimen in every session (s. Fig.7). Confirming this subjective view of the participants, the physical effort of the participants during the training sessions was indeed significantly higher for the robot group in comparison with the control group ($t(14) = 3.170$, $p < .01$, $d = 1.694$, one-sided) (s. Fig.8). Thus, the robot might be a relevant factor to push participants to their optimal physical performance goals. Following this rationale,

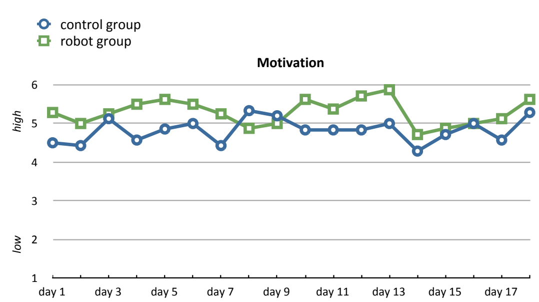


Fig. 8. Subjective Feeling of Motivation

participants in the robot groups felt indeed significantly more motivated during their training than participants in the control condition ($t(14) = 2.086$, $p < .05$, $d = 1.115$, one-sided) (s. Fig.9). This does not s.m to be attributable to simply more fun at the workout as there were no significant differences between both groups regarding training enjoyment (s. above). Hence, the initial analyses suggest that the interaction with the robot led to better training effects, more intensive workouts and also higher training motivation.

XI. Conclusion & Discussion

In this paper, we developed an action-based motivation model for HRI, drawn from qualitative HHI observations. Furthermore, we investigated how this model performed in a HRI setting, in which a robot took the role of a fitness instructor. In order to gain an in-depth understanding of the way in which participants reacted to concrete robot behavior and instructions, we focused on a single participant and presented a qualitative micro-analysis of the video-data of two fragments. The following conclusions can be drawn from the study: The proposed model for motivation with fine-grained and sequential repair strategies has been well received by the user. He appreciated the sequential character of the advices (preparation followed by instruction, repair and finally feedback), allowed the robot time for reacting and thus, was able to put the instruction into beneficial practice. This analysis was supported by a quantitative analysis of the entire sample. The quantitative results suggest that the workout with the robot – compared with the non–responsive control system – led to better training effects, more intensive workout and higher motivation. However, further qualitative as well as quantitative analysis need to be undertaken. As our results suggest long–term interaction between robot and human are characterized by adaptation on the side of the user: participants stop provoking a dialog, do not react upon social triggers (e.g., greeting, etc.) and learn how to operate the robot. Additionally, with repeated interaction users have to cope with faulty reactions and low reactivity. This is why users begin to make strategic use of their interaction. Participants only follow instructions if they consider those as useful. We believe that repair and feedback strategies have been crucial for the increased exercise compliance and that these aspects should feature in all long–term HRI systems.

Appendix: Transcription Conventions

Each tier represents the annotated behavior of instructor (I), trainee (T), robot (R), participant (P) and shows verbal utterances (-ver), gaze (-gaz) or actions (-act). Verbal utterances use the GAT standard, i.e. general spelling in lower case, upper case signifies stressed syllables and punctuation denotes prosodic features (',' = rising; ';' = falling). Important annotation symbols are R = robot , P = participant, @ = at, res = resistance, pedal = cadence, + /- X = in- /decreasing value, Video still are linked to the transcript via their time code.

References

[1] I. Berger, B. Wrede, F. Kummert, and I. Lütkebohle. Towards a robotic sports instructor for high-interaction sports. In *International Conference on Social Robotics, 2013. ICSR 2013.*, pages 589–590, 2013.

[2] T. Bickmore and R. Picard. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2):293–327, 2005.

[3] J. Fasola and M. J. Mataric. Using socially assistive human–robot interaction to motivate physical exercise for older adults. *Proceedings of the IEEE*, 100(8):2512–2526, 2012.

[4] J. Han, M. Jo, V. Jones, and J. H. Jo. Comparative study on the educational use of home robots for children. *Journal of Information Processing Systems*, 4(4):159–168, 2008.

[5] D. Harel. Statecharts: A visual formalism for complex systems. *Science of computer programming*, 8(3):231–274, 1987.

[6] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988.

[7] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro. Interactive robots as social partners and peer tutors for children: A field trial. *Journal of Human Computer Interaction*, 19(1):61–84, 2004.

[8] C. Kidd and C. Breazeal. A robotic weight loss coach. In *Proceedings of the national conference on Artificial Intelligence*, page 1985, 2007.

[9] C. Kidd, D. Cory, and C. Breazeal. Robots at home: Understanding long-term human-robot interaction. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3230–3235, 2008.

[10] S. Kopp, B. Krenn, S. Marsella, A. Marshall, C. Pelachaud, H. Pirker, K. Thórisson, and H. Vilhjálmsson. Towards a common framework for multimodal generation: The behavior markup language. In *Intelligent virtual agents*, pages 205–217, 2006.

[11] D. C. McClelland. *Human Motivation*. Scott, Foresman, 1987.

[12] C. Meyer and U. Wedelstaedt. Körper und ihre individuen: Distributing motivation, koordination und vergemeinschaftung im spitzensport. In *Vielfalt und Zusammenhalt*, number Dokumentationsband zum 36. Kongress der DGS. Deutsche Gesellschaft fÃijr Soziologie, 2014.

[13] C. Midden and J. Ham. Using negative and positive social feedback from a robotic agent to save energy. In *Proceedings of the 4th International Conference on Persuasive Technology*, pages 12:1–12:6, 2009.

[14] J. Peltason and B. Wrede. Modeling human-robot interaction based on generic interaction patterns. In *AAAI Fall Symposium: Dialog with Robots*, 2010.

[15] K. Pitsch and B. Koch. How infants perceive the toy robot pleo. an exploratory case study on infant-robot-interaction. In *Second International Symposium on New Frontiers in Human- Robot-Interaction (AISB)*, pages 80–87, 2010.

[16] K. Pitsch, H. Kuzuoka, Y. Suzuki, P. Luff, C. Heath, K. Yamazaki, A. Yamazaki, and Y. Kuno. "the first five seconds": Contigent stepwise entry as a means to secure sustained engagement in human-robot-interaction. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 985–991, 2009.

[17] A. Powers, S. Kiesler, S. Fussell, and C. Torrey. Comparing a computer agent with a humanoid robot. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, pages 145–152, 2007.

[18] N. Riether, F. Hegel, G. Horstmann, and B. Wrede. Social facilitation with social robots? In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 41–48, 2012.

[19] E. Schegloff. When "others" initiate repair. *Applied Linguistics*, 21(2):205–243, 2000.

[20] G. Schnabel, H.-D. Harre, and J. Krug. *Trainingslehre - Trainingswissenschaft*, chapter 4.1.2, pages 207–211. Meyer & Meyer Verlag, 2008.

[21] S. E. Schultz and D. Schultz. *Psychology and Work Today*. Prentice Hall, 2010.

[22] J. Sidnell and T. Stivers. *The handbook of conversation analysis*. W. John Wiley and Sons, 2012.

[23] L. Süssenbach. Motivation im sport: Eine konversationsanalytische rekonstruktion interaktiver motivationsverfahren im "indoor cycling". Master's thesis, University of Bielefeld, Germany, 2011.

[24] L. Süssenbach and K. Pitsch. Interactional coordination and alignment: Gestures in indoor cycling courses. In *Gesture and Speech in Interaction*, 2011.

[25] L. Süssenbach, K. Pitsch, I. Berger, N. Riether, and F. Kummert. "can you answer questions, flobi?": Interactionally defining a robot's competence as a fitness instructor. In *RO-MAN, 2012 IEEE*, pages 1121–1128, 2012.