Human Perception of Intrinsically Motivated Autonomy in Human-Robot Interaction

Marcus M. Scheunemann[1*] Christoph Salge[1] Daniel Polani[1]
Kerstin Dautenhahn[1,2]

[1]Adaptive Systems Research Group, School of Computer Science,
University of Hertfordshire, Hatfield, United Kingdom

[2]Social and Intelligent Robotics Research Laboratory, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada

February 17, 2020

Abstract

A challenge in using fully autonomous robots in human-robot interaction (HRI) is to design behavior that is engaging enough to encourage voluntary, long-term interaction, yet robust to the perturbations induced by human interaction. Here we evaluate if an intrinsically motivated, physical robot can address this challenge. We use predictive information maximization as an intrinsic motivation, as simulated experiments showed that this leads to playful, exploratory behavior that is robust to changes in the robot's morphology and environment. To the authors' knowledge there are no previous HRI studies that evaluate the effect of intrinsically motivated behavior in robots on the human perception of those robots. The paper discusses challenges of investigating the perception of autonomous, intrinsically motivated robots. We present a game-like study design, which allows us to focus on the interplay between the robot and the human participant. In contrast to a study design where participants order or control a robot to do a specific task, the robot and the human participants in our study design explore their behaviors without knowledge about any specific goals. We conducted a within-subjects study ($N = 24$) were participants interacted with a fully autonomous Sphero BB8 robot with different behavioral regimes: one realizing an adaptive, intrinsically motivated behavior and the other being reactive, but not adaptive. A quantitative analysis of post-interaction questionnaires showed a significantly higher perception ($r = .555, p = .007$) of the dimension "Warmth" compared to the baseline behavior. Warmth is considered a primary dimension for social attitude formation in human cognition. A human perceived as warm (i.e. friendly and trustworthy) experiences more positive social interactions. If future work demonstrates that this transfers to human-robot social cognition, then the generic methods presented here could be used to imbue robots with behavior leading to positive perception by humans.

Keywords intrinsic motivation · autonomous robot · user study · human-robot interaction · predictive information · information theory · embodiment · social robotics

*Contact corresponding author using marcus@mms.ai or visit `https://mms.ai`.

1 Introduction

In this article we look at the effects that intrinsically motivated robot behavior has on the perceptions of robots by humans. This is part of a larger research program to produce autonomous robots, i.e. robots which are not teleoperated or remote controlled, capable of sustained interaction with humans. Sustained interaction should be voluntary, i.e. the human interaction partner should be motivated to interact with the robot without the need for an external reward or without being extrinsically motivated (for more detailed examples see Oudeyer & Kaplan, 2009). This, for example, can be observed in a child interacting with a puppy. The child will likely be motivated to do so, even without an external reward (such as promised money) and even without the existence of an extrinsic reward (such as playing with the puppy as a means to an end, i.e. to train it). Instead the motivation for the interaction might results purely from wanting to do this activity for its own sake, i.e., the child is intrinsically motivated to play with the puppy. This effect should also be extended over time, beyond the novelty effect, i.e. not wear off at all, or at least not quickly. To stay with the example with the puppy, the first encounter might be very exciting and engaging. This excitement may decrease, but there is something in the interaction which often keeps children engaged over a longer period of time. In long-term HRI this observation is explained with the novelty effect. The novelty of interacting with a robot is motivating for humans, but wears off relatively quickly to the point that they show no wish for further interactions (Dautenhahn, 2004; Leite et al., 2013).

There are approaches to counteract the novelty effect. For example, Pinillos et al. (2016) developed an autonomous hotel robot. It attracts attention by the hotel guests, many of them wanting to know more about the robot itself. They propose that the robot's services (i.e. its competence or usefulness) needs to be large in order to keep customers engaged. On the other hand, Kanda et al. (2010) developed a semi-teleoperated mall robot and incrementally added novel behaviors, such as self-disclosure. A field trial indicates that the robot attracted reoccurring visitors, without increasing its services. Engagement is also a concern in the field of social robotics in education (Belpaeme et al., 2018). One existing approach here is to develop robots with a set of hand-designed questions, comments and statements (Gordon et al., 2015; Ceha et al., 2019). This makes the robots appear curious, which elicits curiosity in the humans too, which in turn enhances learning and memory retention (Oudeyer et al., 2016). Curiosity is part of the broader concept of intrinsic motivation (Oudeyer et al., 2016), or is even used synonymously for intrinsic motivation (Schmidhuber, 1991). It should be noted that, in contrast to the work proposed by Schmidhuber (1991), the previously mentioned robots were programmed with behavior to have them *appear* curious. Their behavior was not actually generated by some curiosity formalism. The studies either needed a constraint context, a specific task (e.g. Gordon et al., 2015; Pinillos et al., 2016), or were relying on humans teleoperating the robot (e.g. Kanda et al., 2010; Ceha et al., 2019). Teleoperation, or the Wizard-of-Oz model, remains the state of the art for many HRI studies (Clabaugh & Matarić, 2019). This is caused by the challenge to define a sufficient set of execution rules (i.e. behaviors) for an HRI task; this holds true even in a laboratory setting. It remains elusive to achieve an autonomous, social behavior in an unconstrained environment, i.e., for any given task or goal in the real world (Christensen et al., 2016; Belpaeme et al., 2018). Developing a robot driven by an actual intrinsic motivation formalism, such as the drive to explore its environment and its capabilities, might offer a solution to both problems. If successful, this would provide us with a robust behavior generation mechanism that allows us to "Escape Oz" (Clabaugh & Matarić, 2019), while also producing behavior that appears curious, or similarly engaging to the human interaction partner. This will reduce the reliance on human adaptation or teleoperation, and could provide a promising pathway towards having robots more easily deployed in the every day life.

Our idea is that imbuing a robot with a computational model of intrinsic motivation (IM) make it perceived as a genuine *social other* – similar maybe to an animal – and thus be of more interest to a human interaction partner. The concept of intrinsic motivation originates in psychology, initially in close relation to Self-Determination Theory (SDT) (Ryan & Deci, 2000b). SDT posits that humans have an inherent tendency to seek out novelty and challenges, to extend and exercise their capacities to explore and to learn, without having to be coerced by an extrinsic reward. According to SDT, humans have inherent drives for competence, autonomy and relatedness. Computational models of intrinsic motivation aim to formalize the principles that create those drives to make them operational, i.e. they can be used to create spontaneous exploration and curiosity in an artificial agent (Oudeyer & Kaplan, 2009). Therefore, we hypothesize that they will give an artificial agent a stronger social presence, and thus make them a more interesting interaction partner.

The known models of intrinsic motivation have a range of interesting properties. The idea of universality is of particular interest for this application, in particular the fact that IMs can cope with changes to an agent's environment or its morphology. This makes this approach, in principle, suitable to be deployed on any robot and it also allows it to deal with any environment or context. The biggest limitation here is usually computational complexity. The method is also limited by the fact that several approaches at least require agent-centric forward models, similar to sensorimotor contingencies (O'Regan & Noë, 2001), which might not be easily obtainable. Finally, most IMs can be expressed to operate on the immediate perception-action loop of the robot, allowing for tightly coupled or entrained behavior with both the environment or other actors. Both of these properties make IMs an interesting family of approaches to deploy in autonomous human-robot interaction (HRI) robots, as there is a requirement for interactive feedback on a short feedback loop and for the ability to robustly deal with a range of situations. This is particularly relevant, as social cognition is believed to heavily depend on interaction – and thus any approach that aims to encourage interaction, should be robust to the perturbations induced by those social, and possibly physical interactions.

In the remainder of this paper we want to substantiate this main idea with an extensive HRI study involving 24 human participants. As a first step we evaluate the human perception of robots with different behavior with the help of post-experiment questionnaires. We compare how the introduction of intrinsically motivated behavior affects human perception, and discuss how these factors can lead to formation of different social attitudes. Our main focus in this paper is on the dimension of perceived *Warmth*[1]. Warmth and *Competence* are considered the two main dimensions in describing almost all social attitudes in human social cognition, such as, e.g., friendliness, empathy, admiration, envy, contempt and pity (Fiske et al., 2007; Abele et al., 2016). Warmth is considered the primary dimension for social characterizing peers. This means, when characterizing other people, we firstly judge their intent (Warmth) before judging their capability (Competence) to enact their intent. Warmth is strongly linked to the measure of trust (Fiske et al., 2007; Fiske, 2018). A person who is perceived as warm is also perceived as more trustworthy. For example, Kulms & Kopp (2018) use it as an indicator for people's trust in computers. From social cognition, it is known that human's who are perceived as warm experience more social interaction than their peers who are perceived less warm. Consequently, in order to welcome robots in our everyday life, an understanding is needed for how to enable the perception of Warmth for robots.

We will see that the intrinsically motivated robot is perceived as more warm than a baseline robot it is compared with. This is a step towards the long-term goal of producing a robot capable

[1]We will continue to capitalize dimensions like Warmth to indicate that we refer to, e.g., the questionnaire dimension of Warmth, as opposed to true perceived warmth. At times, however, we will use the adjectives (e.g, warm) where it is clear that we refer to the dimension.

of sustained interaction, as it suggests a method to induce a positive social attitude towards the robot in the human. Further studies are, of course, warranted to see if this effect transfers from human-human interaction to human-robot interaction. In other words, we still need to investigate if higher perceived Warmth for a robot actually leads to more sustained interaction. The interplay between personality and social relationships is still an ongoing – and complex – investigation for human-human interaction (Geukes et al., 2019). Our working hypothesis is that a robot which is perceived as warm (i.e. friendly and trustworthy) is more likely to receive more positive, longitudinal interactions.

The paper is structured as follows. First, we will outline some background on intrinsic motivation, its computational approaches and its relation to autonomy, insofar it relates to the present work. We will then specifically introduce predictive information (PI), the formalism we use to implement intrinsic motivation in our studies. We will outline the concrete approximations (and their assumptions) to compute PI. Our description will, in particular, highlight how to make this approach suitable for deployment on an actual robot and why it is a good candidate for our research questions.

We then present two within-subjects studies: a preliminary study ($N = 16$) and the main study of this paper ($N = 24$). Both studies consists of two conditions with the same robot platform: the behavior in one condition is generated using predictive information, the behavior in the other condition is a reactive baseline behavior. The focus of all conditions is the interplay between the robot and the human participant. Importantly, in contrast to many HRI studies, participants cannot order the robot to do something. Instead, the robot and the human participants explore their behavior towards each other. The preliminary study has been conducted and published prior to this work (Scheunemann et al., 2019). We will summarize the preliminary study, to introduce the first steps towards investigating human-perception of an intrinsically motivated robot, and to present design decisions, like for example the baseline behavior, which is used in both studies. The preliminary study didn't significant effect. However, it indicated that an intrinsically motivated robot may be perceived as more warm than the reactive baseline behavior, whereas the reactive baseline behavior may be perceived as more intelligent.

We then present the main study, which has been designed according to the learned lessons from the preliminary study. We present its design and the results, concentrating on our two main hypotheses: one focusing on the perceived Warmth of the intrinsically motivated robot, and the other on the lack of difference in Perceived Intelligence or Competence between the robot scenarios. We found that our game-like study design makes both robot behaviors appear similarly competent. This is important in order to focus solely on the Warmth dimensions, without interfering by the Competence rating. Most importantly, the main study provides evidence that an intrinsically motivated robot is perceived as more warm compared to a robot using the reactive baseline behavior. We will discuss the implications of those findings, and how they can be applied to other projects.

2 Background

In this section provides some background of the previously mentioned concepts relating to this work.

2.1 Intrinsic Motivation

A common definition of intrinsic motivation (IM) in psychology is "doing [...] an activity for its inherent satisfactions rather than for some separable consequence" (Ryan & Deci, 2000a). An intrinsically motivated agent is moved to do something, not for a separable consequence, but due to an inherent drive. Since intrinsic motivations have been considered an instrumental ingredient in the development of humans (Oudeyer et al., 2007), there has also been a great interest in developmental robotics to produce formalized models that can be used to imbue robots with drives for competence and knowledge acquisition (Oudeyer & Kaplan, 2008, 2009). Nowadays, there is a range of formal models that roughly fall under the header of intrinsic motivation, such as the free energy principle (Friston, 2010), predictive information (Ay et al., 2008), homeokinesis (Der & Martius, 2012), empowerment (Klyubin et al., 2005), learning progress (Kaplan & Oudeyer, 2004), the autotelic principle (Steels, 2004), and others. These models have a range of commonalities: they are free of semantics, task-independent, universal and can be computed from an agent's subjective perspective. Most of the work related to IMs focuses on how they create "reasonable" behavior (in some suitable sense) for simulated agents. There has been some work in the domain of computer games that focuses more explicitly on the relationship between intrinsically motivated agents and humans, and how an intrinsic motivation could generate more believable Non-Player Characters (NPCs) (Merrick & Maher, 2009), or produce generic companions (Guckelsberger et al., 2016) or antagonist behavior (Guckelsberger et al., 2018). So far, IMs have been deployed on simulated and physical robots (e.g. Oudeyer et al., 2007; Der & Martius, 2012; Martius et al., 2014), but, as far as we know, there has been no human-robot interaction study yet evaluating the perception of intrinsically motivated robots from the perspective of humans.

2.2 Autonomy

The term *autonomy* is overloaded (Boden, 2008) and used with different meanings in this paper. When we talk about autonomous robots, we merely mean robots that are not directly controlled by a human operator, autonomy just being a dimension of the experimental design (Huang et al., 2004; Stubbs et al., 2007). In self-determination theory (SDT), however, autonomy refers to being in control of one's own life, which can be seen as a close enough analogy for living systems (Paolo, 2004). SDT also assumes that there is a drive to maintain this state of autonomy, which we do not see in general with autonomous robots. We might see autonomy used as the idea that a robot should strive to maintain operational autonomy, i.e. not be in need of external help, but it usually does not refer to a robot striving to not be controlled by a human. Finally, autonomy might also be referring to the concept of self-making or self-law-giving, which is closely related to autopoesis (Maturana & Varela, 1991; Froese & Ziemke, 2009). In robots, this is currently only a theoretical idea (Smithers, 1997), but it is often considered necessary for "true" intrinsic motivation. Any heteronomy during the development or creation of an agent would ultimately make them extrinsic and hence undermine their very nature, i.e. computational models of intrinsic motivations on robots are usually put on those robots by humans, and are thus actually extrinsic. Computational models of intrinsic motivation are an attempt to merely reproduce the behavior or functionality of genuinely intrinsic motivation in organism. This is also the reason that we talk about *perceived* agency and *perceived* autonomy. One idea behind this is that by using those models for the robots to "pretend" to be intrinsically motivated, humans might indeed perceive the robot as thus. In the following, when we talk about intrinsic motivation on the robot we will exclusively refer to the initial, technical meaning, the computational model that aims to mimic intrinsic motivation. While the more philosophical underpinnings of autonomy are highly relevant to the larger context of this work and will make this approach useful even if we develop robots with more extensive autonomy, we will set them aside for the present work.

3 Predictive Information as Candidate for Intrinsic Motivation

This section describes predictive information (PI), the intrinsic motivation model used for the robot behavior generation in our experiments. PI has been described as early as 1986, termed *effective measure complexity* (Grassberger, 1986) or *excess entropy* (Crutchfield & Young, 1989). Previous work with PI-driven robots in simulation demonstrated its applicability to a large range of different robot morphologies (Der et al., 2008; Martius et al., 2013b; Zahedi et al., 2013; Martius et al., 2014). A range of existing videos[2] showcase apparent exploratory, playful and open-ended behavior of individual robots and robot collectives. The PI-induced behavior in the videos suggests PI as a promising immediate candidate measure to test our core idea.

Conceptually, when this measure is transformed into a behavior-generating rule, the resulting dynamics essentially falls into a family of learning rules related to the reduction of the time prediction error in the perception-action loop of a robot (see especially the book "The Playful Machine", Der & Martius, 2012). The aforementioned book also shows how these approaches can be computed from the robot's perspective alone. Additionally, the variety of different robots and their behaviors presented there shows how different behaviors arise from the same formalism due to the sensitivity towards the agent's specific embodiment.

The predictive information formalism consists in computing a specific learning rule that aims to maximize the mutual information between a robot's past and future sensor states (Ay et al., 2008), i.e., PI quantifies how much information a history of past sensor states contain about future sensor states. More generally, predictive information is defined as the mutual information between the past and the future of a robots sensory input. A high amount of predictive information requires two things: First, past sensor states should make future sensor states more predictable. This should lead the robot to act so that its actions have predictable consequences. Furthermore, the robot also needs to create a high variety of sensor input. If the robot always would perceive the same sensor input, then there is either insufficient information in the past to predict future sensoric states, or an insufficiently varied future for which there is not much to predict. In both cases, an impoverished sensoric input reduces the predictive information. Alternatively, if there is strong variation in sensoric inputs but little structure in the sensory data stream, i.e. the past has little to do with the future, that will also lead to low predictive information. Vice versa, a high value for predictive information requires a high entropy in future sensor states, i.e. a richly varied future (a robot motivated to "excite" its sensors to reach a rich variety of different states) which at the same time depends on the observable past (i.e. which the robot can predict well based on the past). The behavioral regime is created by these two counterpoised requirements: predictability and variety. This yields a robot wanting to act so that its future is highly predictable, while exploring and experiencing new sensor states. The PI literature argues that this balancing act produces rich exploratory behavior that is sensitive to the robot's embodiment and argues that predictive information is "the most natural complexity measure for time series" (Bialek et al., 2001; Martius et al., 2013b).

Der et al. (2008); Ay et al. (2008, 2012) presented derivation rules for PI, which allows for computing the model directly for linear systems with stationary dynamics. The next subsections will present an extension of their work by Martius et al. (2013b) for the use in nonlinear and nonstationary systems – such as physical robotic systems. The main idea is that instead of computing the full system dynamics, only the system's time-local dynamics are considered to compute the PI values. This quantity is called time-local predictive information (TiPI) and is the one used in this work. Subsection 3.1 will provide an overview of TiPI and introduce into

[2] Video page, including works with PI: https://robot.informatik.uni-leipzig.de/videos.

the measure. This is followed by a section presenting the derivation of the explicit update rules used for our studies (3.2). The deviations are kept short to provide the basic concepts of the quantity and introduce the underlying main approximations and assumptions. Subsection 3.3 will discuss these approximations and assumptions with respect to applying TiPI in an HRI scenario. Subsection 3.4 will summarize this section.

3.1 Overview of time-local predictive information

The predictive information formalism to generate the robot's intrinsically motivated behavior in the studies of this article is closely following the implementation of Martius, Der, & Ay (2013b). They propose an approximation to compute PI for nonlinear systems with nonstationary dynamics, which allows for behavior development of a self-determined robotic system. They approximate PI with assuming small, Gaussian noise and only consider a time window over the current state of the robot and τ steps back in the past, coining it time-local predictive information (TiPI). TiPI allows for going beyond discrete finite-state actions, which still dominates scenarios of information theory-based behavior generation, towards continuous actions. This permits the use in physical robots in high-dimensional state-action spaces. TiPI enables robot behavior with self-switching dynamics in a simple hysteresis system and spontaneous cooperation of physical coupled systems Martius et al. (2013b).

It works by updating the two internal neural networks of the robot, one that generates behavior from sensor input and the other that predicts the future states. The continuous adaptation, aimed at improving the TiPI, moves the robot through a range of behavioral regimes. Importantly, the changes in behavior are partially triggered by the interaction with the environment, as mediated through the robot's embodiment. The rate at which those internal neural networks are updated is the one model parameter which could be adapted for individual preferences (Der & Martius, 2006).

The approach allows to change the robot's morphology without having to redesign the algorithm, but will still remain sensitive to the embodiment of the robot, meaning that the resulting behavior differs, depending on how the robot interacts with the world. The morphology can be changed by changing physical parts or by choosing different sensors as inputs for the robot's neural networks. In both ways, the robot can be guided towards exploring and playing in different ways. For example, by including a sensor for the robot's angular velocity around its main axis, the spherical robot will try to spin clockwise and anticlockwise with changing velocities. If we further include an accelerometer providing measurements of the forward and backward acceleration, the robot will try to explore the relationship between spinning movements and locomotion, yielding a variety of additional motion patterns. If, furthermore, a human is interacting with the robot, this can increase the behavioral diversity, depending on the interaction between the robot and the human.

3.2 Deriving update rules for time-local predictive information

Martius et al. (2013b) present estimates of the time-local predictive information (TiPI) for general stochastic dynamical systems. For systems with Gaussian noise and with gradient ascent on the TiPI landscape, they derive explicit expressions for exploratory dynamics. We do not aim to provide a full mathematical background of the method. For a detailed treatment, the reader should refer to (Ay et al., 2008; Martius et al., 2013b).

Assume a robot has n sensors and the sensor readings are polled in constant time steps ($\Delta t = 1$). Combine now the result of all sensor values in a vector $s \in \mathbb{R}^n$. A series of those sensor readings between points of time a and b (with $a < b$) can be described as a time-discrete process $\{S_t\}_{t=a}^{b}$, where both boundaries are included. Let the past be defined by the points of time $a, \ldots, t-1$ and the future by $t, \ldots, b$. Bialek et al. (2001) defines the PI for some point in time t for the time series S as the mutual information between the past and the future. Intuitively, the mutual information measures the shared information of two random variables, here S_{past} and S_{future}, i.e., it measures how much knowledge of the past S_{past} reduces the uncertainty of the future S_{future}. The predictive information, expressed as mutual information, is thus defined as follows

$$\begin{aligned} I(S_{\text{future}}; S_{\text{past}}) &= \left\langle \ln \frac{p(s_{\text{future}}, s_{\text{past}})}{p(s_{\text{future}})p(s_{\text{past}})} \right\rangle \\ &= H(S_{\text{future}}) - H(S_{\text{future}}|S_{\text{past}}) \end{aligned} \tag{1}$$

with the average taken over the joint probability density distribution $p(s_{\text{past}}, s_{\text{future}})$.

The first essential simplification proposed by Martius et al. (2013b) is applying the Markov assumption to Equation 1. If $\{S_t\}_{t=a}^{b}$ is a Markov process, all past information relevant to the future is stored in the very last state of the system, i.e. $S_{\text{past}} = S_{t-1}$.

The predictive information in this case reduces to:

$$\begin{aligned} I(S_t; S_{t-1}) &= \sum_{s_{t-1} \in S_{t-1}} \sum_{s_t \in S_t} p(s_t, s_{t-1}) \ln \left(\frac{p(s_t, s_{t-1})}{p(s_t)p(s_{t-1})} \right) \\ &= H(S_t) - H(S_t|S_{t-1}) \; . \end{aligned} \tag{2}$$

In general, the Markov assumption will only hold true for real-world sensor processes in exceptional cases. Nonetheless, as in the wide use of e.g. particle or Kalman filters, it is a popular assumption for successfully approximating problems using a Bayesian approach (Thrun et al., 2005). Martius et al. (2013b) use the reduced Equation 2 as the definition of the objective function for deriving the autonomous exploration dynamics.

Above Equation 2 is a quantity derived for the whole process. However, to create an actual behavior rule that reacts to current situation, it necessary to compute a local quantity, specific to the current situation. Therefore, instead of computing the probability distribution $p(s_t)$ over the whole process, we additionally condition the PI on a state s_{t-2}. The new quantity derived is then

$$I(S_t; S_{t-1}|s_{t-2}) \tag{3}$$

Because of above Markovianity, this is effectively a time-local quantity for PI and therefore it is called *time-local* predictive information (TiPI). To calculate the TiPI, a model of S_t needs to be learned to predict its time series. Let $\psi = \mathbb{R}^n \to \mathbb{R}^n$ be a function predicting the time series at $t-2$, $t-1$ and t via

$$\hat{s}_{t-2} = s_{t-2} \tag{4}$$
$$\hat{s}_{t-1} = \psi(s_{t-2}, \theta_{t-2}) \tag{5}$$
$$\hat{s}_t = \psi(\psi(s_{t-2}, \theta_{t-2}), \theta_{t-1}) \tag{6}$$

In an example implementation[3], ψ is realized as a one-layer neural network. θ is a set of

[3]Online: `https://github.com/georgmartius/lpzrobots/blob/d2e6bbd164d902cdaa57eef154ed353ee0027236/selforg/controller/pimax.cpp`.

parameters representing the synaptic weights and biases, which will be updated each time step in order to increase TiPI. The actual dynamics of the process can described via

$$s_t = \psi(s_{t-1}, \theta_{t-1}) + \xi_t \tag{7}$$

ξ_t being the prediction error.

We denote the deviation of the actual dynamics (Equation 7) from the deterministic prediction (Equation 6) as

$$\delta s_{t'} = s_{t'} - \hat{s}_{t'} \tag{8}$$

for any time t' with $t-2 \le t' \le t$. Since s_{t-2} is the initial state for TiPI, there is no deviation at time $t-2$ and $\delta s_{t-2} = 0$, while one step after the initial state $\delta s_{t-1} = \xi_{t-1}$. Intuitively, δs_t represents the prediction error(s) accumulated from the start of the prediction (here at $t-2$) up to time t.

For very small prediction errors the dynamics of δs (Equation 8) can be linearized as an approximation:

$$\delta s_{t'} = L(s_{t'-1})\delta s_{t'-1} + \xi_{t'} + O(||\xi_t||^2) \tag{9}$$

with the Jacobian

$$L_{ij}(s) = \frac{\partial \psi_i(s,\theta)}{\partial s_j}$$

Assuming that the prediction errors ξ are both small and Gaussian, the TiPI on the deviation process $\delta S_{t'}$ is the same as on the original process S_t (see Martius et al., 2013a, sec. A). It is therefore sufficient to concentrate on the error propagation for the computation of the TiPI. This reduces Equation 2 in such a way that only the probability distribution of the deviation $p(\delta s)$ needs to be known, rather than the probability distribution over the full state $p(s)$.

If we further assume that the prediction error ξ is white Gaussian, the entropy can be expressed as covariances (Cover & Thomas, 2012). The resulting explicit expression of TiPI on δS becomes:

$$I(\delta S_t; \delta S_{t-1}|s_{t-2}) = \frac{1}{2}\ln|\Sigma_t| - \frac{1}{2}\ln|D_t| \tag{10}$$

with $\Sigma = \langle \delta s\, \delta s^T \rangle$ as the covariance matrix of the process δS, and $D = \langle \xi \xi^T \rangle$ as the covariance matrix of the prediction error. Note that the predictive information becomes meaningful only at t, as the prediction error vanishes at $t-2$ and at $t-1$ the two covariance matrices coincide: $\Sigma_{t-1} = D_{t-1}$. The covariances are exact for Gaussianity. For the general case they are approximations only.

We now give the algorithm used to drive a robot's behavior towards increasing TiPI. Martius et al. (2013b) derive it explicitly for the gradient ascending neural network presented in Equation 6. They argue that the prediction error ξ is essentially noise and does not depend on the parameter of the controller, and that therefore the term $\ln|D|$ of Equation 10 can be omitted when computing the gradient. Based on Equation 10, the resulting gradient step executed at each time t is

$$\Delta\theta_t = \epsilon\frac{\partial I}{\partial \theta} = \epsilon\frac{\partial}{\partial \theta}\ln|\Sigma_t| \tag{11}$$

with ϵ being the update rate and $\theta_{t+1} = \theta_t + \Delta\theta_t$.

Applying Equation 9 to above equations results in explicit gradient step

$$\Delta\theta = \epsilon \left\langle \delta u_t^T \, \frac{\partial L(s_{t-1})}{\partial \theta} \, \delta s_{t-1} \right\rangle \tag{12}$$

where δs and the auxiliary δu are given as

$$\begin{aligned}
\delta s_{t-1} &= s_{t-1} - \psi(s_{t-2}, \theta_{t-2}) \\
\delta s_t &= s_t - \psi(\psi(s_{t-2}, \theta_{t-2}), \theta_{t-1}) \\
\delta u &= \Sigma_t^{-1} \delta s_t \\
\Sigma_t &= \langle \delta s_t \, \delta s_t^T \rangle
\end{aligned}$$

To render $\Delta\theta$ computable the Equation 12 is further approximated by applying the *self-averaging property* (this will be explained in more detail below) of a stochastic gradient

$$\Delta\theta = \epsilon \, \delta u_t^T \, \frac{\partial L(s_{t-1})}{\partial \theta} \, \delta s_{t-1} \tag{13}$$

As per (Der et al., 2008; Martius et al., 2013b), Equation 13 is the equation by which the (approximate) TiPI maximization is ultimately implemented. We remark that increasing $|\Sigma|$ corresponds to an increase of the norm of δs. In other words, this reflects the amplification of small fluctuations in the motor dynamics, i.e. an increase of the instability of the system dynamics.

3.3 Considerations with applying TiPI to the real world

Martius et al. (2013b) apply the above maximization of TiPI to simulated robots. As a result, those robots show complex behavior. One example is a humanoid robot with 17 degrees of freedom (DoF) controlled by a single high-dimensional controller implementing the PI optimization principle from Equation 13. Importantly, despite using the same rules, the formalism produces different behavioral regimes of the simulated humanoid, depending on the environment it is exposed to. Along the above derivation, several approximations and assumptions have been made. When the measure is applied to a real robot in a real-world human-interaction scenario, this requires a careful considertion of the assumptions and approximations, which we do in the following.

3.3.1 Markov assumption

This assumption simplifies the definition of the objective function Equation 2. More importantly, it renders TiPI (Equation 3) computable as it simplifies the conditional probability density distribution. Applying the assumption to robotics-related problems, especially to make Bayesian problems manageable, is common in robotics (Thrun et al., 2005). This approximation therefore can be considered a popular robotics strategy for applying information theory and Bayesian algorithms to the real world.

3.3.2 Conditioning on an initial from two states back

To compute PI for nonlinear systems with nonstationary dynamics, the proposed solution is to condition the quantity on an initial state being two steps back in time. We stick here to the minimal possible window mainly because computing a larger window online comes to a computational cost challenging to bear on embedded systems.

The sensors used for the input need to be meaningful for the time window. For example, a global position of the robot does not change much within the time window of two steps, so the robot cannot excite he sensor value in the chosen window. This reduces the choice of sensors to the ones displaying variation within the given time window, It is therefore preferable to choose sensors which display variation within the given time window, such as proprioceptive sensors measuring the acceleration or velocity.

3.3.3 Prediction errors are both: very small and Gaussian

These assumptions are made at various places for deriving the explicit update rules. For example, the assumptions were used to show that TiPI on the process δS (propagation of errors) is equivalent to the one on the original process S (sensor states). This enables the linearization of the error dynamics Equation 9 and eventually, under the same assumptions, the formulation of explicit TiPI expressions (Equation 10). Assuming that the error is very small and Gaussian has implications on choosing the right sensors for the experiments. Therefore, care needs to be taken that the noise of the sensors remains somewhat Gaussian and somewhat small for the duration of the time window. For example, the motor position typically changes in a continuous fashion and therefore the respective sensors will be good candidates to fulfill this assumptions.

On the contrary, it would violate the Gaussianity assumption to use a sensor whose values exhibit, e.g. sudden drops, such as proximity sensors based on Bluetooth (Scheunemann et al., 2016). Such sensors measure the signal strength to an external device which is prone to occlusions and can sometimes intermittently fail to provide any reading at all. To mitigate this, it is possible to use filters to smoothen the sensor readings.

3.3.4 Applying the self-averaging property for stochastic gradients

Equation 13 uses the so called self-averaging property of stochastic gradients, that means, that a stochastic gradient over larger number of steps in a sequence acts as an approximation of averaging over the probability distribution (Van Rensburg et al., 2001). In other words, we can replace the average over multiple independently drawn samples by a one-shot gradient.

Practically, this makes Equation 10 computable, as the density distribution of the gradient is hard to obtain. Martius et al. (2013b) note that using this property is only exactly valid for a small update rate ϵ, when it is driven to zero eventually. Note that the update rate ϵ in our application is quite large to allow for a very fast adaptation process. Martius et al. (2013b) argue that the explicit update rules favors the approach of getting an "intrinsic mechanisms for the self-determined and self-directed exploration", with the exploration being driven only by the sensor values. Thus, the one-shot nature of the gradients favors the explorative nature of the exploration dynamics and increases interesting synergy effects, but is not strictly implementing the average.

3.3.5 Noise is independent of the controller parameters

To derive the explicit update rules (Equation 10), the covariance of the noise $D = \langle \xi \xi^T \rangle$ is omitted altogether. The propagation in error is only assumed to be pure independent noise in the environment. In other words, the noise is independent of the controller parameter θ. Martius et al. (2013b) justify this because of the "parsimonious control" implemented by the formalism.

All these assumptions are of course no longer strictly valid once the robot interacts with the environment, especially human. Nevertheless, the intended richness of the robot's behavior is not hampered by that. Instead, the formalism gives rise to a varied and manifold repertoire of behaviors, as shown by many studies mentioned in (Ay et al., 2008; Der & Martius, 2012; Martius et al., 2013b, 2014).

3.4 Summary

The TiPI method generates aforementioned variety of different behavioral patterns for a robot. This makes TiPI-maximization a promising candidate for use in HRI settings. Its universality for different embodiments and nonstationary settings makes it a good candidate for applying it to a robot without concerning oneself too much with the environment or the robot's particular embodiment. Completely missing from the existing body of work on TiPI, however, is the actual evaluation of the behavior when it is induced by the interaction with humans. This is the gap this paper aims to fill.

4 Robot & Measures

This section describes the robot and the used evaluation measures, both used in a preliminary study (Scheunemann et al., 2019, described in the next section) and the follow-up study, the main study of this article, described afterwards.

4.1 Robot

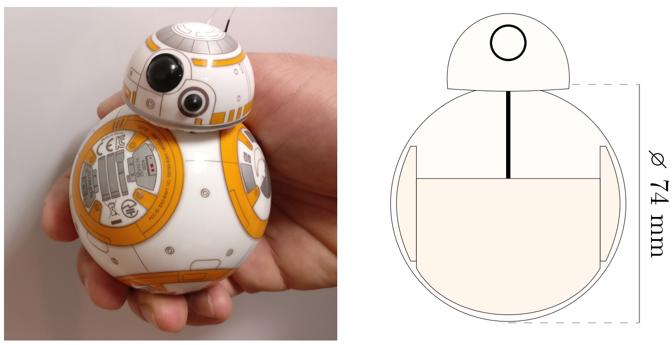

Figure 1: Left: The used robot platform BB8 from Sphero. Right: A 2-D cross-sectional view of the robot. A two-wheel vehicle (darker shape), kept in position by a heavy weight, moves the sphere when driving. The speed of each servo motor can be set individually, allowing the robot to move straight, to turn and to spin. A magnet attached to the vehicle keeps the head on top of the sphere facing in moving direction.

It is argued that an anthropomorphic robot platform may raise expectations of its social capabilities in participants (Dautenhahn, 2004; Hayashi et al., 2010), which in turn may interfere with the investigation of the perception of intrinsic motivation. We therefore decided to use a more simplistic platform with a few degrees of freedom, to maximize the focus on the effects induced by TiPI solely. We decided for the off-the-shelf spherical robot from the company Sphero[4], specifically, the BB8 platform, as depicted in Figure 1. BB8 is a character from the "Star Wars"

[4]Information about the company and its products: https://support.sphero.com.

movies[5]. A magnet keeps the head in driving direction, which gives the user a sense of the robot's direction. We think that this helps the human participants to interact with the robot.

The robot's on-board hardware is proprietary in such a way that we could not flash it or run our own code. However, it is possible to communicate with the robot using Bluetooth Low Energy (BLE). We can, for example, request a stream of sensor information from the robot, or control the robot by either using (1) the robot's balance controller or (2) directly setting the speed of each servo.

The robot's balance controller is an in-built controller for locomotion. We refer to this controller as a "balancing" controller and will denote conditions where the robot is using this controller with a subscript b. For example, the condition with an intrinsically motivated, adaptive robot directly setting its servo speed will be denoted as ADA. If the intrinsically motivated, adaptive robot is using the balancing controller, the condition is denoted as ADA_b.

The balancing controller receives speed and heading as input values. The heading is globally initialized to zero degrees when the robot is started. This means, if you send 20° to the controller, the robot will always set its heading towards 20 degrees on top (clockwise turn) of the initial heading. This controller is a closed-loop controller. If the robot gets nudged or turned, it will try to keep the previously set heading constant.

It is also possible to directly set the speed of the left and right servo. This is an open-loop controller[6], always setting the speed without any further observations.

As for sensors, the robot offers raw sensor information from a 3-axis accelerometer, a 3-axis gyrometer and the actual motor speed of each servo measured as voltage of the *back electromotive force* (back EMF). The robot can stream data from an inertial measurement unit (IMU) represented in quaternions or Euler angles[7]. Additionally, it offers velocity information along a plane in the x and y direction, and also positional data (i.e. odometry) estimated from its starting position. For our studies, we use sensor data from the IMU, the accelerometer, the gyrometer and the speed of the wheels.

4.2 Measures

Up to now, all relevant quantities were objective quantities that can be obtained from the sensors used by the given robot. We now proceed to characterize its counterpart, the human partner, in the dynamics. We use two standardized scales to measure the participant's perceptions of the robots: the Godspeed scale (Bartneck et al., 2009), which has been widely used in many HRI experiments, and the Robotic Social Attributes Scale (RoSAS), a more recently designed questionnaire by Carpinella et al. (2017), which has seen relatively little use in HRI to date[8].

Godspeed uses a 5-point semantic differential scale and investigates the dimensions *Anthropomorphism*, *Animacy*, *Likeability*, *Perceived Intelligence* and *Perceived Safety*. The RoSAS tests for the dimensions *Warmth*, *Competence* and *Discomfort*. Carpinella et al. (2017) do not recommend a specific size for the Likert-questions, but recommend including a neutral value, e.g., by having an uneven number of possible responses. Our questionnaire consists of 7-point Likert-type items. Note that we will continue to capitalize the dimensions to indicate that we refer to, e.g.,

[5] Online: https://www.starwars.com.

[6] Note that when we say open-loop, we mean it only from the robot's perspective, as the robot firmware applies the data to the robot hardware in an open-loop fashion. From the perspective of the overall PI-maximizing behavior, of course, the data sent to the robot indeed depends on earlier sensor values.

[7] Note that this sensor is faulty along the roll axis for when the robot is rolled for more than 90 degrees. This is, however, not an issue as the sensor is only used as an input for the robot when using the balancing controller, and extreme roll cases ($|\alpha_{\mathrm{roll}}| > 90°$) are not reached.

[8] Examples of both questionnaire can be found here https://gitlab.com/scheunemann/latex-questionnaire.

the questionnaire dimension of Competence, as opposed to true competence. At times, however, we will use the adjectives, e.g, competent, where it is clear that we refer to the dimension.

The dimensions Warmth and Competence are central dimensions for evaluating other humans as social beings. According to Fiske et al. (2007), people perceived as warm and competent elicit uniformly positive emotions, are in general more favored, and experience more positive interaction by their peers. The opposite is true for people scoring low on these dimensions, meaning they experience more negative interactions. Warmth and Competence, together, almost entirely account for how people perceive and characterize others.

Grossly simplified, perceived Warmth leads to generally positive or negative social bias, referred to as active facilitation (Cuddy et al., 2007). High perceived Warmth usually results in a positive bias in human social cognition. The Competence dimension mostly moderates this effect. High Warmth and high Competence result in admiration, while high Warmth and low Competence result in pity (Judd et al., 2005). The corresponding effects for low Warmth are envy and contempt. As a result, Warmth can be considered the primary factor for predicting the valence of interpersonal judgments (Fiske et al., 2007; Abele et al., 2016). This means, it primarily predicts whether an impression is positive or negative.

5 Preliminary Study and its Implications

We conducted a preliminary study with 16 participants[9], prior to the work presented here. The findings in that preliminary study are the basis for the current study. We think that describing the preliminary study has two advantages. Firstly and most importantly, it helps to understand the current study design, e.g., the chosen baseline behavior, the human-robot interaction tool and the game-like scenario. Secondly, we think that because of the novelty of the investigation, it may prevent future studies from making similar non-optimal design choices. The next sections are brief descriptions of the study design and its results and implications for the current study. More details on the preliminary study are discussed in (Scheunemann et al., 2019).

In the preliminary study we mostly focused on the perceived Animacy and Intelligence of a robot. Contrary to our expectations, we found that an intrinsically motivated robot, i.e., a robot using predictive information maximization for behavior generation, as described in section 3, is not perceived as more animated, intelligent or competent than a baseline behavior. Instead, we found that the intrinsically motivated robot is perceived as more warm (i.e. friendly or trustworthy).

5.1 Preliminary Study: Description

Our aim was to investigate whether an intrinsically motivated robot is perceived as more animated (i.e. animal-like) or competent/intelligent compared to a baseline behavior. This novel investigation opened several questions: (1) What is a fair baseline behavior for comparison? (2) How should we design the study so that people are encouraged to interact with a robot in a way where they can see behavioral differences? (3) How can we measure Animacy or Competence? This section answers these questions in that order.

[9] 5 females and 11 males with an average age of 33.4 years with $SD = 9.3$. Detailed sample information in (Scheunemann et al., 2019, sec. 3.8)

Baseline Behavior

We considered the following three alternative means of behavior-generation for serving as a baseline:

1. human remotely controls the robot
2. random behavior
3. pre-adapted reactive behavior

Ideally, we want to see how the intrinsically motivated robot (using TiPI maximization) compares to a human remotely controlling the robot. However, human controlled behavior has a high degree of variance, dependent on the particular human controller, making it challenging to use as a reproducible baseline behavior across different studies and different research groups. Furthermore, it is unclear how much access the human controller should have to environmental information. If the human can directly observe the participants, they could obtain much more information than a robot which has only its built-in sensors, giving them an unfair advantage in creating behavior responsive to the participant. If we limit the human controller to only the robots' sensors, then the human controller would likely struggle to make sense of this input, and potentially even be unable to control the robot at all.

The problem with using random behavior as a baseline is that "randomness" actually has a set of parameters that needs to be chosen, such as how often particular values will change, or also whether it is the *change* of the value or the value itself that is being randomized. We performed some preliminary trials with random values, but were quickly facing the question of a fair baseline behavior again. Having the experimenter choose these values leads to basically designing a certain kind of behavior (chosen from a whole range of behaviors). Again, this raises the issue of reproducing a fair baseline comparison over different studies, as the experimenter influences the behavior with the chosen parameters.

We decided to use a pre-adapted reactive behavior as it offered the highest degree of similarity to the intrinsically motivated, adaptive behavior that we would like to compare it with. The pre-adaptation is done with the very same PI implementation and parameters, using the same sensors as the robot will use in the preliminary study. We conducted three 5 min runs with the PI implementation. We then froze and stored the network configurations of the robots, and chose one of them randomly as the starting configurations for all trials of the experiment. The resulting parameters are dependent on the robot morphology and the environment, rather than on the experimenter. The resulting behavior is reactive, as the sensor value will in general change the robot's behavior. However, the robot is not adaptive in this phase as the sensor weights do no longer change.

Preliminary Study Design

Figure 2 shows the experimental environment. Two tables form an area of 180 cm $\times$ 120 cm in size that the robot can move around in. The surface of the table differs in friction and height. The black foam area has a hill (top) and a pit (bottom). Additionally, the black area and the white paper area are softer and have higher friction compared to the wooden part.

The participants' task is to observe the robot and understand whether it has a strategy for exploring the environment. The participants were told "Please find out whether the robot follows a specific strategy for exploration". The robot arena is open to one side where the participant is supposed to stand and interact with the robot, as shown in Figure 2. The participant has an additional task, which is to keep the robot from falling over the edge. The participants were told:

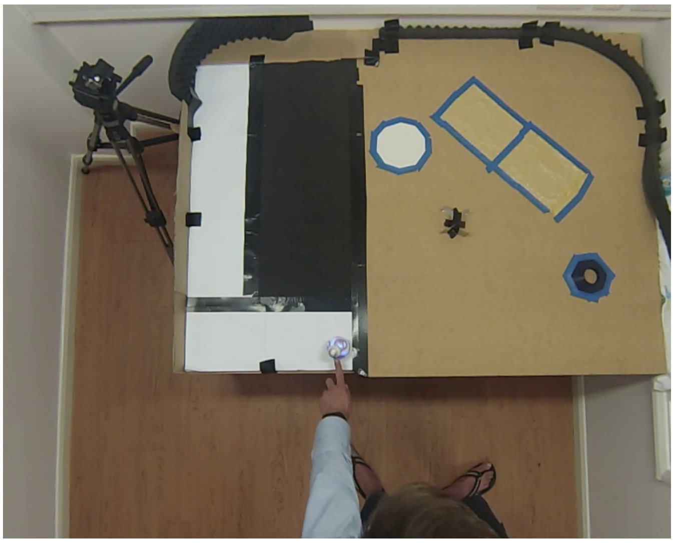

Figure 2: The environment the robot explores during the trials from a birds eye perspective. The white area is paper, the black is foam material and the beige colored area is wood. At the top of the foam material is a hill area and a pit in the lower part. The bottom edge does not have a wall, forcing the participant to interact with the robot in order to prevent it from falling off the table.

"Please keep the robot from falling over the table edge. You can either use your flat hand to block it [Investigator shows the motion] or you can nudge the robot to keep it away [Investigator shows the motion]. To understand whether the robots have a different strategy, you can also interact with the robot at any time". The hope was that these instructions will encourage the participants to interact with the robot, and that this enforcement of interaction would provide the participants with a better understanding of the robot's capabilities and behavioral richness. However, as we will discuss later, we observed that the participants assumed an implicit goal of the robot not falling off the table.

The experiment consists of two different conditions of 10 min each:

- REA_b (reactive): participants interact for approximately 10 min with a reactive robot (i.e. the baseline behavior) and are asked about what they have seen.
- ADA_b (adaptive): same as REA_b, but the robot is continuously adapting, based on maximization of predictive information, as a motivation to interact with its environment.

The order of REA_b and ADA_b is randomly assigned, but counterbalanced over the number of participants. Table 1 shows the group label and the condition order of both groups, along with the number of participants.

Table 1: Groups and their order of conditions in the preliminary study.

group	order of conditions	participants
A	$ADA_b \rightarrow REA_b$	8
B	$REA_b \rightarrow ADA_b$	8

There are 5 inputs to the robot's networks: the pitch and roll angle from its IMU; the linear velocity from the forward/backward axis and the linear velocity from the left/right axis from the accelerometer. The final input is the angular velocity around the upright axis of the robot. The input sensors were chosen empirically, but in close consideration of the dynamics of TiPI, presented in section 3. In both conditions, the robot uses its balanced closed-loop controller. This means the controller receives heading and speed information, and applies it to the robot in such a way that the robot aims to be upright while following the commands. We found that, if the robot then tries to excite its roll and pitch angle, some interesting behavior emerges. We also decided to input the linear forward acceleration and the angular velocity, which, in a way, maps to the controller input of setting its heading and its speed. Exciting those sensors can be done by changing the controller values directly. Both conditions have the same sensory input to allow for a fair baseline comparison. This means that both conditions will be reactive to the sensor input, but only the ADA_b condition will continue updating the network weights and biases, whereas the REA_b condition has those weights and biases constant over the experiment.

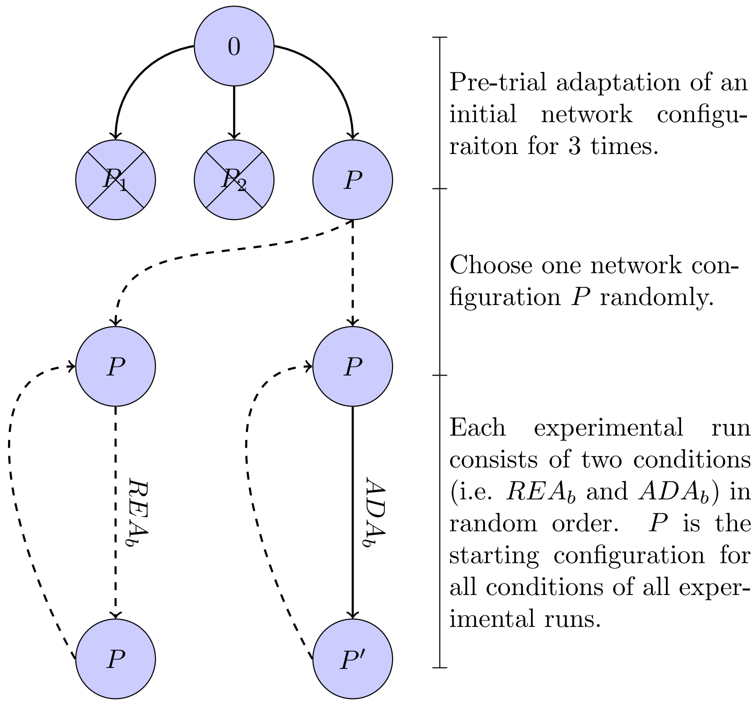

Figure 3: The starting point of each condition is a network configuration P. The network configuration P is chosen randomly from three 5 minute pre-adaptation trials. Only in the ADA_b condition the adaptation is continuous during the human-robot interaction trial. In condition REA_b, the robot is only reactive and not adaptive towards the environment.

Figure 3 shows how the starting configuration is derived for all networks of both conditions. They are generated in two steps. Firstly, three trials with the robot for 5 min in the previously described environment are conducted. At the end of each trial, the robot's network configurations are saved. In a second step, one of these network configurations is randomly chosen as the starting configuration, i.e., for condition ADA_b and REA_b.

The TiPI formalism allows for having different levels of adaptivity to changing environments and new stimuli. The update rate for ADA_b was determined empirically as follows. We noticed that the robot can get caught in the pit mentioned earlier. If that happens, the robot needs to adapt to leave the pit and continue exploration. The ADA_b adaptation rate was set so that the robot would change its behavior and leave the pit in less than 20 s. We hypothesized that a high adaptation rate yields a higher perceived intelligence, as the robot would continuously adapt to new stimuli and change the way it reacts to the environment, i.e., new inputs. The robot in the ADA_b condition was therefore assumed to be perceived as more intelligent, as it will be the only one able to leave the pit. Furthermore, we hypothesized the intrinsically motivated, adaptive robot will be perceived as more animal-like, as it is adapting to the sensor input by the human. Example videos for both conditions are available from (Scheunemann, 2019).

Measures

After each condition, a participant is given a questionnaire encompassing the standardized scales: the RoSAS and the Godspeed scale, as described in subsection 4.2. We used them as a tool to understand whether the intrinsically motivated, adaptive robot is perceived as more animated/animal-like or competent/intelligent compared to the reactive baseline behavior.

After the last condition, i.e., after both conditions have been conducted, participants were asked to complete two additional open-ended questions:

1. "Can you describe the different behaviors of the robot? Did the robot have any particular strategy for exploring?" and
2. "What were the best and/or worst aspects of the robots behavior?".

This questions helped to understand whether participants see any differences in the robot behaviors.

5.2 Results and Implications for the current Study

Table 2: The results as presented in (Scheunemann et al. (2019)). Reported are the p values and the confidence intervals of a Wilcoxon Signed Rank Test of all dimensions of the RoSAS and Godspeed scale comparing between REA_b and ADA_b. The standardized effect size r indicates that there is a medium effect for Warmth and Perceived Intelligence, a large effect for Discomfort, and small effects for Animacy, Likeability and Perceived safety.

	dimension	95% confidence interval lower bound	95% confidence interval upper bound	p	r
RoSAS	Warmth	-0.67	0.17	.37	.32
RoSAS	Competence	-0.58	0.50	.80	.09
RoSAS	Discomfort	-0.83	0.08	.14	.52
Godspeed	Anthropomorphism	-0.30	0.40	.92	.04
Godspeed	Animacy	-0.25	0.33	.70	.14
Godspeed	Likeability	-0.30	0.40	.73	.12
Godspeed	Perceived Intelligence	-0.20	0.80	.24	.41
Godspeed	Perceived Safety	-0.67	0.67	.44	.27

Table 2 shows detailed results as presented in (Scheunemann et al., 2019). The participants could distinguish the robot behavior between the two conditions. However, the adaptive robot was not perceived as more intelligent or competent than the reactive baseline robot. In fact, there was rather evidence for the opposite: we found a medium effect that the reactive robot was perceived as more intelligent than the adaptive robot. We did not expect that, as for us, the intrinsically motivated robot shows adaptation to the environment making it more intelligent than the only reactive robot. However, from the perspective of the participants, this may just be blurred as the intrinsically motivated robot is more adaptive, but its behavior switches turned out to make the robot approach the edges of the table more often. This in turn felt not intelligent for some participants. This might be due to our instruction about keeping the robot from falling off. In hindsight, this might have given some participants the belief that an implicit goal of the robot is to not fall of the table – and consequently they might have seen the behavior of moving towards the edge as a cognitive failing, rather than an exploration.

We were thinking that a different introduction to that additional task could make a difference. For example, the participants could be asked "if the robot is seeking attention and approaches you, try to not let it fall over the edge". Either way, both scenarios will bias participants in some manner.

We therefore decided to change the study design for the current study. Our initial approach was to enforce the interaction between the human and the robot, by giving the participant an assisting task of preventing the robot from rolling over the edge. This time we designed a more game-like set-up, where the participant is not assisting the robot in its own goal, but the participant is tasked with figuring out, through interaction, if the robot has the same or different behavior. While this tasks still focuses the participant on the robot behavior, it is less biased in regards to suggesting an implicit goal for the robot. It should also change the initiative to the participant, as they would not have to react to the robot, but could engage and interact as they choose.

We found that the intrinsically motivated robot is perceived as more warm than the baseline behavior. Warmth is an important dimension to explain social attitudes in social cognition (as discussed before) and the finding therefore caught our attention. However, to confirm this finding, we consequently formulated a follow-up hypothesis for testing in our main study. Based on the observed effect size it also seemed prudent to increase the number of participants to increase the statistical power of our analysis.

Result from our preliminary study suggest a medium effect that participants perceive the reactive robot as more intelligent than the intrinsically motivated robot. There is evidence that the perception of, e.g., Competence, influences the dimension Warmth (Carpinella et al., 2017). Due to this interference we think the best case for our follow up study would be a design where we see no effect for Competence or Perceived Intelligence, so we can focus on testing for Warmth. With our modifications to the participants instructions, and based on our results from the preliminary study, we now hypothesize that we will not see an effect for Competence or Perceived Intelligence.

Because this is still one of the first HRI studies into the human perception of intrinsically motivated robot, we decided to measure and report on all dimensions of the RoSAS and Godspeed questionnaires. Not only could this help to potentially identify further minor effects (as happened for us in the preliminary study), but it will also provide other researchers attempting similar investigations a baseline for comparison.

6 Main Study

This section describes the study design, the sections also point out the differences to the preliminary study where needed. The robot platform and the used questionnaires are the same for the preliminary study and the current study and have been described above in section 4.

6.1 Robot, Environment and Tasks

Figure 4: The picture shows the author using the interaction tool. He nudged the robot with the white end of the wand. Participants were able to freely chose a position around the table for observing or interacting with the robot.

The robot will locomote on the table shown in Figure 4. It is circular, with 91 cm in diameter and 27 cm in height. A foam wall of 2.5 cm in height and with 4 cm in width surrounds the border of the table. We decided for these measurements in such a way that the robot cannot fall off the table, even with a very high velocity. Three blankets of a total height of 3 to 4 mm cover the surface (including the walls). This applies some friction and makes it easier for the robot to locomote on the otherwise smooth and slippery surface of the wooden table top. The table's distance to the surrounding wall of the room is at least 60 cm, allowing participants to freely move around the table.

Figure 4 shows the first author of this article interacting with the robot using the HRI tool referred to as a *wand* that was developed specifically for this study. Participants were asked to use that and touch and nudge the robot with the white end. The wand is 50 cm long and weighs 78 g. It consists of a 40 cm long aluminum tube with a diameter of 10 mm. The end is a round, softer sphere. It is made of a an off-the-shelf table tennis ball with a diameter of 40 mm.

There are two major differences to the preliminary study described in section 5. Firstly, all borders of the interaction environment are enclosed. If the participant decides to be passive, the robot cannot fall off the table. Also, the round shape of the table and its position allows participants to reach all borders.

The other major environmental difference is the interaction itself. The participants were asked to use the wand for interacting. We assumed that this will help to ease the interaction, as some participants in the preliminary study felt uncomfortable with the idea of using their hands for means of interactions.

6.2 Sample

We recruited 24 participants (10 female; 14 male) mostly from university staff and students, between the ages of 18 and 64 years ($M = 31.7, SD = 12.6$). The participants were undergrads or post-graduates from the university, but all naïve towards the goal of the experiment. 8 participants have a background in HRI, whereas 9 participants never participated in any prior HRI study. All were asked how familiar they are with interacting with robots, programming robots and the chosen robot platform. 5-point Likert-questions were chosen with the value 1 for “not familiar” to 5 for “very familiar”. The self-assessed experience for interacting with robots was an average of 3.5 (Mode = 5). The average familiarity with programming robots was 3.2 (Mode = 5) and the experience with the chosen robot platform was rated an average of 2.1 (Mode = 1). The familiarity with the movie series “Star Wars” was rated an average of 3.2 (Mode = 4).

The study was conducted on the premises of the University of Hertfordshire and was ethically approved by the Health, Science, Engineering & Technology ECDA with protocol number aCOM/PGR/UH/03018(3). The anonymity and confidentiality of the individual data is guaranteed.

6.3 Condition and groups

This experiment consists of two conditions with the following characteristics:

- REA_b **(balanced, reactive):** The robot uses its balanced mode for locomotion, the network controlling the robot has been pre-adapted using PI and it remains constant. The binary running the robot and the weights of the network are exactly the same as in the preliminary study (cf. 5). The name of the condition is therefore kept the same.
- ADA **(unbalanced and directly controlled, adaptive):** The robot is not using its balanced mode, but rather controls both servos directly. The robot is adapting continuously, based on maximization of TiPI.

The reactive robot in the REA_b condition starts with the same networks as the robots in the preliminary study. The weights are received based on pre-trial adaptation. This determines how it reacts to sensor input, but it does not further update its internal network during the experiment. The reason for taking the REA_b robot from the first study are two-fold. Firstly, the behavior is a good baseline behavior. The robot was interesting to the participants and the behavior was not too simple so that the participants did not see any patterns. Secondly, keeping the baseline constant, but changing other variables, allows for a better comparison to the previous findings and the previous adaptive robot.

The intrinsically motivated robot in the ADA condition realizes behavior motivated by TiPI maximization, and it continuously updates its internal networks based on that gradient during the experiment. In contrast to the preliminary study, the robot controls its two servos directly. This allows for a generation of robot behavior of more variety, i.e., using a larger variety of servo configurations. The robot sensor input is again the linear acceleration for the forward/backward and left/right axis from the accelerometer, and the angular velocity around the upright axis received by the gyrometer. Instead of using the absolute position of the robot received via its pitch and roll angles from the IMU as in the preliminary study, we now input the speed of the two servos. This allows for a direct coupling between the output of the controller changing the servo speed, and the actual measured servo speed.

We wanted the robot to behave similarly in the beginning as the robot in the REA_b condition. Therefore, we tweaked the starting weights of the network by hand. As there is direct coupling between the servo speed readings and the controller output, i.e., the set speed for the servos, the weights were set in such a way that a reading on the left servo would amplify the output for the left servo, and vice versa. This way we could create a slow-pace forward movement for the first few seconds, which looks similar to the reactive baseline robot.

Table 3: Groups and their order

group	order of conditions	participants
A	$ADA \rightarrow REA_b$	12
B	$REA_b \rightarrow ADA$	12

The condition is an independent within-subject variable, meaning all participants are exposed to both conditions. The conditions are presented in a randomized, but counterbalanced order to the participants. This order splits the participants into two groups displayed in Table 3.

The participants interact with the robot for 5 min in each condition and they will be given a questionnaire after each interaction. Details will be discussed in subsection 6.6.

6.4 Research questions

The study is concerned about two research questions. Firstly, we want to understand whether the changes in the study design is more "fair". We know that the perception of Competence influences the perception of Warmth, our first research question is therefore:

- Does a more game-like scenario help to make the two robot behaviors appear similarly competent or intelligent? In other words, does a game-like scenario and the existing baseline behavior truly concentrate on the behavioral level, without blurring the results through an assumed goal-directedness for one of the robots?

We hypothesize that the answer is "yes". This means, for the dimensions Competence and Perceived Intelligence we would expect to see no evidence for an effect.

In the preliminary study we saw evidence that participants perceive an intrinsically motivated robot as more warm. This study shall confirm whether the observed effect is existing for a larger sample size:

- Is an intrinsically motivated, adaptive robot perceived as more warm than a reactive baseline behavior?

Again, we hypothesize that we will find evidence that this is the case. We do not expect any other strong effects, but we will report and discuss the main effects of all other dimensions.

6.5 Measures

Similar to the preliminary study (cf. 4.2), two standardized scales are used to investigate for participant's perception of the different robot behaviors. The RoSAS and the Godspeed scale are used in the post questionnaires, i.e., the questionnaires handed to the participant after each of the both conditions REA_b and ADA.

6.6 Procedure

Participants are welcomed to the experimental room, are then handed an information sheet and are asked to sign an informed consent form. Then the environment and the robot is presented and briefly described.

Other than in the preliminary study, participants are not given a specific robot's aim or are asked to prevent the robot from following over the edge. On the contrary, they are only asked that their task is to observe whether the two presented robots are different. For seeing differences, they can use the HRI tool: the wand. They are allowed to nudge the robot or block it. Both of these actions are demonstrated to the participants. However, no other information is provided.

The idea is that participants are not expecting the robot to perform some task at a particular level and thus concentrate on that aspect in detail. In the preliminary study, for example, having a task to prevent the robot from rolling over the edge, made participants aware of that and it seems that most judgment about the robot's capability was whether the robot can avoid the edge or not. In the current study, the scenario is much more game-like and less directed, and thus allows us to focus on the interplay between the robot and the human. The participants, as described, are provided with a tool to interact with the robot and their only goal is to find out if the robots are different.

Participants then complete a pre-questionnaire. This gathers information regarding their gender, age and background. Next, the two conditions are presented to the participants in a randomized order, each lasting approximately 5 min. They fill in a post-questionnaire containing the two scales after each condition. The entire experiment takes 50 to 60 min per participant.

6.7 Data preparation

To prepare the data for analyzing we tested the score reliability of the scales of both standardized questionnaires with the use of Cronbach's α. We found that the item *quiescent-surprised* is negatively loaded on the dimension Perceived Safety. Even if reversed, the reliability is poor with $\alpha = .54$. Therefore, we removed that item. Table 4 presents all test results, revealing a good score reliability ranging from .75 to .85 and acceptable reliability for the dimension Anthropomorphism: $\alpha = .67$.

Table 4: Internal consistency reliability measured with Cronbach's α

	dimension	items	α
RoSAS	Warmth	6	.80
	Competence	6	.85
	Discomfort	6	.80
Godspeed	Anthropomorphism	5	.67
	Animacy	6	.75
	Likeability	5	.82
	Perceived Intelligence	5	.82
	Perceived Safety	2	.82

7 Results

We analyzed the responses to the questionnaires with the use of non-parametric tests. We will first analyze the answers for possible interaction effects, i.e., does it make a difference how the conditions are presented to the participants (7.1). We then present the main effects of each dimension (7.2).

7.1 Interaction effects

Table 5: ANOVA-type test results for the independent variables "order" (levels: A, B), "condition" (levels: REA_b, ADA) and the interaction of both variables "order:condition" for the dimensions of the RoSAS and Godspeed scale.

		order			condition			order:condition		
	dimension	F	$df1$	p	F	$df1$	p	F	$df1$	p
RoSAS	Warmth	0.098	1	.755	11.733	1	<.001	0.001	1	.976
	Competence	0.163	1	.687	0.047	1	.828	1.473	1	.225
	Discomfort	1.365	1	.243	1.143	1	.285	1.787	1	.181
Godspeed	Anthropomorphism	0.278	1	.598	13.810	1	<.001	0.100	1	.751
	Animacy	0.931	1	.335	14.789	1	<.001	0.634	1	.426
	Likeability	0.786	1	.375	1.856	1	.173	1.455	1	.228
	Perceived Intelligence	0.424	1	.515	0.001	1	.977	0.031	1	.860
	Perceived Safety	10.477	1	.001	1.343	1	.246	1.381	1	.240

An analysis of variances (ANOVA) is commonly used for investigating for interaction effects, i.e., effects that show that the order of the conditions influences the responses of participants to a condition. Because of the relatively small sample size ($N = 24$) we decided for using a non-parametric ANOVA-type test.

The study has two independent variables: one within-subjects variable and one between-subjects variable. The within-subjects variable, i.e., the independent variable that all participants are exposed to, is the condition. It consists of the two levels REA_b and ADA (cf. 6.3). The between-subjects variable is the independent variable unique to each participant. This variable is the order of the both conditions.

Table 5 shows the results of a non-parametric ANOVA-type test[10]. The last column "order:condition" reveals the likability for an interaction between the conditions and their order. None of the p values (sub-column p) is smaller than .05. This means for a 5% significance level there is no statistical significance and there is not enough evidence for an interaction effect for any of the dimensions. This is particularly true when looking at the dimension Warmth, the p value is the largest and almost equals to one. This means the presence of an interaction effect is highly unlikely. Without any presence of an interaction effect, we can safely investigate the main effects independently of their order, i.e., we can compare the responses to both conditions independently of whether the participants were exposed to, e.g. ADA, in the beginning of the experiment or at the end.

[10]For computing the ANOVA-type test we used the `R` package `nparLD`. As the study consists of one within-subjects variable (condition) and one between-subjects variable (order), it can be expressed as F1-LD-F1 Model. The `nparLD` package offers the function `f1.ld.f1()` for computing such models.

7.2 Main effects

Looking at the second column "condition" of Table 5 and its three sub columns, there is evidence of statistically significant effects for the dimensions Warmth, Anthropomorphism and Animacy. To understand the direction, i.e., is the perceived Warmth higher for the condition ADA or the condition REA_b, we use a paired difference test. The Wilcoxon signed-rank test is a potential candidate known to be robust for small sample sizes. We used the two-sided version of the test to investigate for effects in both directions. We report the test statistic V, the p value, a point estimate and its according confidence interval. The point estimate (short: estimate) is the median of the difference between REA_b and ADA. It provides a size and a direction for how much the participants prefer one condition. For example, if the estimate between REA_b and ADA equals -0.833, this means that on average the participants responded to Warmth with 0.833 units higher in the ADA than in REA_b. The units here are the responses to the Likert-type items ranging from 1 to 7 (RoSAS) or the differential scale ranging from 1 to 5 (Godspeed). Along with the estimate we further report the standardized effect size r. It allows for investigating the size of a potential effect independently of the sample size. Cohen (1992) defined the effect as small when $r > .1$, as medium when $r > .3$ and as large when $r > .5$.

Table 6: Main effects for all dimensions of the RoSAS and Godspeed scale for comparing REA_b and ADA. The p value shows statistical significance (*) for the dimensions Warmth, Antrophomophism, Animacy and Likeability. This provides evidence that each participant responded differently on those dimensions for each of the conditions REA_b and ADA. The standardized effect size r indicates that the effect is medium for Likeability and large for the other dimensions.

			95% confidence interval				
	dimension	V	lower bound	upper bound	p	estimate	r
RoSAS	Warmth	27.5	-1.333	-0.333	*.007	-0.833	.555
	Competence	138.5	-0.833	0.583	.988	0.000	.003
	Discomfort	54.0	-1.250	0.417	.287	-0.250	.217
Godspeed	Anthropomorphism	25.0	-1.300	-0.500	*.002	-0.900	.643
	Animacy	30.5	-1.250	-0.417	*.002	-0.833	.636
	Likeability	38.0	-0.700	-0.100	*.038	-0.400	.424
	Perceived Intelligence	144.5	-0.500	0.500	.875	0.000	.032
	Perceived Safety	49.0	-0.500	1.250	.422	0.500	.164

Table 6 shows the results of the two-sided Wilcoxon signed-rank test[11] for all dimensions of both standardized scales comparing the condition REA_b and ADA. We can see that a large and statistical significant effect ($r = .555, p = .007$) for the dimension Warmth. The estimate is negative. This is because, on average, participants respond higher to the robot in the ADA condition, making the difference of $REA_b - ADA$ negative. In other words, most participants perceived the robot in the ADA condition as more warm than the robot in the REA_b condition. This directly answers our second research question (cf. 6.4), namely that an intrinsically motivated robot (as the one in the ADA condition) is perceived as more warm.

Figure 5 visualizes the magnitude of the effect. The magnitude of the effect increases with an increasing distance of the estimate to zero. It also visualizes the certainty that the point estimate is indeed the true effect. The smaller the error bars, i.e., the confidence interval, the

[11] R's `wilcox.test()` is used.

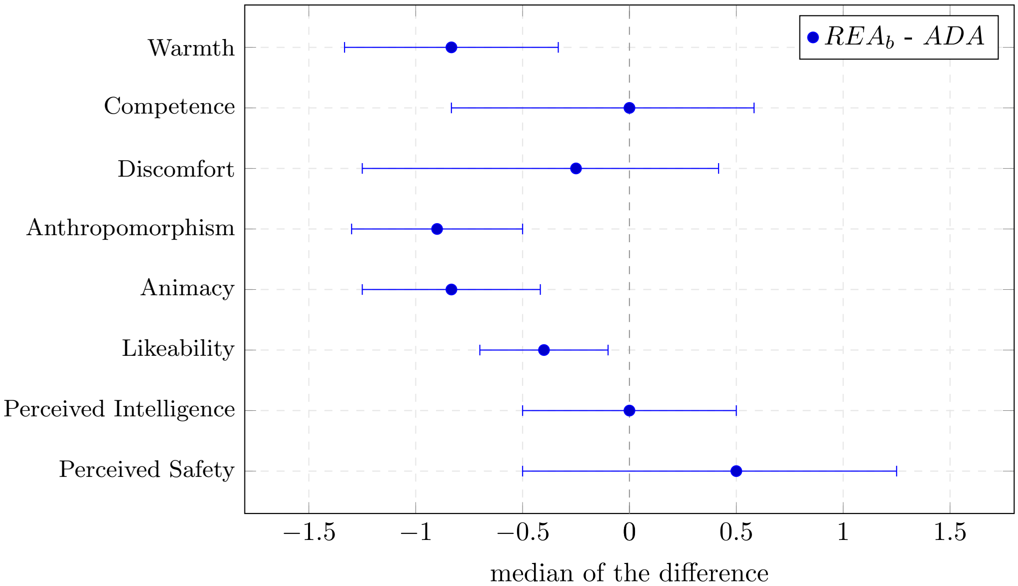

Figure 5: The results of the two-sided Wilcoxon signed rank test: a paired difference test and a non-parametric alternative for the paired t-test. The median of the difference is plotted as point estimates and the 95% confidence interval as error bars. The smaller the error bars, the more certain is the effect. The bigger the distance of the point estimate to 0, the larger is the effect. The graph also helps for visualizing statistical significance. If 0 is not included in the error bars, a statistical significance is present, i.e., $p < .5$.

more certain we can be about the point estimate. Figure 5 confirms that there is a large effect for Warmth in favor of the *ADA* condition.

An even larger effect can be observed for the two dimensions Anthropomorphism and Animacy. For both dimensions, their perception differs and is statistically significant. The estimate and $r > .5$ again indicates that there is a large effect in favor of the *ADA* condition. The p value for the dimension Likeability is statistically significant ($p = .038$) and the effect size $r = .424$ (medium). There is also a small effect for the dimension of Discomfort ($r = .217$), although the robot in *ADA* is perceived as more warm. It feels contradicting at first, but participants can respond high for Warmth and Discomfort at the same time (Carpinella et al., 2017). We indeed saw this already in the preliminary study in section 5.

As hypothesized, there is no statistical significance for neither of the two dimensions Perceived Intelligence and Competence, more importantly, there is no magnitude of an effect. Figure 5 shows that the estimate is close to zero for both of the dimensions. The confidence interval is almost equally distributed around zero and is quite large, indicating that there is no certainty for an effect in any direction. This helps to answer our first research question (cf. 6.4), namely whether a game-like scenario helps to make the two robot behaviors appear similarly competent or intelligent.

8 Discussion

The study results provide evidence that the intrinsically motivated robot is perceived as more warm than the reactive baseline robot. This is an indicator that a PI-driven behavior may prove relevant for human-robot interaction, as the dimension Warmth is one of the universal dimensions for humans judging social attributes on other humans (Judd et al., 2005; Fiske et al., 2007; Abele et al., 2016; Fiske, 2018). Notably, a high scoring for Warmth is considered positive, i.e. desirable. The results leave no doubt that the intrinsically motivated robot is perceived more positively than the reactive baseline robot.

The changes undertaken in the current study design towards a game-like scenario helped us to focus on the Warmth dimension. Neither the Competence nor the Perceived Intelligence dimension scored high for any of the conditions. This is evidence that the participants did not know if the robot had any goals. Although Competence and Warmth are mainly considered unique dimensions, some interference and correlations have been pointed out between them (Fiske et al., 2007; Abele et al., 2016). The lack of an effect for neither Perceived Intelligence nor Competence is therefore an important feature of our study design, which allows for an isolated observation of the influences of IM on the perception of Warmth.

We also – unexpectedly – observed that participants perceived the intrinsically motivated robot as more animated (and anthropomorphised). There is evidence that humans perceive a robot higher in Animacy when the robot moves more “naturally” (Castro-González et al., 2016). In fact, any object is considered animated if it changes speed and direction without visible influences (Tremoulet & Feldman, 2000). Another influence of the perception of Animacy is the reactivity of the robot (Fukuda & Ueda, 2010). We designed our baseline behavior to provide both: similar movement variety, and reaction to sensor input, to allow for a fair comparison and focus on the effects of the intrinsic motivation. In the current study, the control mechanism for the IM robot was changed (cf. 6.3). Other than the reactive baseline behavior, where the robot could only move forward and was kept mostly upright due to the balancing controller, the IM robot had a different behavioral regime. It could go backward and forward, and because the servo speed was set directly and individually, it could produce different behavioral regimes like, e.g., a wobbling locomotion. Therefore there are three explanations for the baseline behavior being perceived as less animated: (1) the motion patterns, (2) the reactivity, or (3) the intrinsic motivation. With the current data, we cannot answer this question sufficiently, but we tend to be skeptical and don’t want to argue for 3 before carefully observing the baseline behavior. Although the baseline behavior showed feasible in the preliminary study, we will therefore investigate possible changes for a follow-up study, which makes the baseline similarly perceived animated as the intrinsically motivated robot here.

However, it needs to be noted that the robot with the baseline behavior is not perceived as inanimate. Instead, participants simply perceived the intrinsically motivated robot as more animate as the baseline. Although this is an indication for the baseline behavior to have less natural movements (as discussed), there is no evidence in the literature that the rating for Warmth has been significantly influenced. In the preliminary study, for example, participants did perceive the baseline behavior more animate (small effect), but they perceived the intrinsically motivated robot as more warm (medium effect). Given the results of both studies, we argue that there is evidence that the different participant responses for Warmth between the two behavior conditions is mainly caused by the robot’s intrinsic motivation.

What remains unclear is how much the knowledge from social cognition transfers to human-robot interaction. Despite recent advances (e.g. Mieczkowski et al., 2019), future work has to understand whether the concepts from social cognition transfer to *physical* interaction with a robot. If that is the case, our study shows that a robot which has intrinsic motivation can help to increase the interest of a human to interact with it, and that an intrinsically motivated robot is likely to perceive more positive social interactions.

Recent research has shown that more complex, anthropomorphic robot platforms mimicking being intrinsically motivated, can help to engage humans (Gordon et al., 2015; Ceha et al., 2019).

To close the gap between the robots mimicking IM, i.e. with behavior implemented and designed by humans, and intrinsically motivated robots presented here, some tasks need to be addressed. They are ranging from more pragmatic straightforward task to quite fundamental questions: For example, can the presented implementation be computed on a more complex robot? Or: how can a robot conceptualize and store explored behavior regimes, and how can it memorize them in similar contexts?

Despite those long-term challenges, our findings offer some direct applications to more current robots. The here presented TiPI formalism can be used to implement a generic motion for times when a robot does not play a specific behavior, i.e. no human is interacting with the robot. We provided evidence that this may attract more humans due to their perception of the robot being friendly (Warmth). This would reduce the times researchers need to hand-tweak natural or affective behavior. To make a robot more engaging and to elicit curiosity in the human interaction partners, we discussed that some researchers proposed that novel behaviors, or a larger variety, are important. These behaviors (questions and statements) are often randomly chosen in autonomous robots (e.g. Gordon et al., 2015). We propose that a more naturalistic selection could be applied with using an intrinsic motivation measure. Using TiPI directly is not the best candidate, as questions or statements cannot be represented by a continuous variable. However, TiPI could be used as a reward signal for a selection algorithm based on, e.g., reinforcement learning. Alternatively, researchers could decide for another formalism implementing IM, like for example empowerment (Klyubin et al., 2005). We argue that our presented study design can help to prototype an affective behavior, or affective behavior selection, for a variety of IM formalism in a relatively short time.

9 Conclusion

We started this research with the question if intrinsically motivated autonomous robots can be beneficial for designing engaging human-robot interaction (HRI). We conducted a within-subjects study ($N = 24$) where participants interacted with a fully autonomous Sphero BB8 robot with two conditions with different behavioral regimes: one realizing an adaptive, intrinsically motivated behavior and the other being reactive, but not adaptive. We used time-local predictive information (TiPI) maximization as one candidate measure to produce intrinsic motivation (IM)-based behavior, and produced, to our knowledge, the first study quantitatively relating human perception of intrinsically motivated robots. Of particular interest is the high similarity between both conditions in Perceived Intelligence ($r = .032, p = .875$) and Competence ($r = .003, p = .988$), which gives support to our non-task oriented interaction design. This was particularly important as Competence ratings can influence the perception of Warmth, which is the dimension we focused on in our study.

Our main result is that the perception of Warmth by human participants is high for the adaptive, intrinsically motivated robot ($r = .555, p = .007$). This is in comparison to a baseline behavior that includes both: similar movement and reaction to sensor inputs – meaning that the difference in perception arises from the robot's adaptation to the physical interaction. This effect was also robust to physical interaction, i.e. it was present even though the robot was physically nudged by the human participants. The dimension of Warmth is, as mentioned previously, an important factor for attitude formation in human-human social cognition. However, it is not immediately clear if this higher perceived Warmth leads to a positive attitude or preferences in human-*robot* interaction. If future work would demonstrate this, then we believe the formalism presented here could be utilized to create a preference or positive attitude towards a robot in a large range of HRI scenarios.

The open questions going forward are now: Can we confirm the results with using a baseline behavior with more similar motion regimes to further strengthen the focus on the intrinsic motivation (IM) of the agent and the interaction? Does the universal applicability of the formalism also translate into a universal, or at least widespread, evocation of Warmth across different robot morphologies? Does this effect persist over time? And does a positive social attitude lead to more engagement? All these questions are empirically testable, and given the positive results here are possible directions for future research.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

MS, CS and KD conceived of the presented idea. MS worked on the theory and implemented and tested it on the robot. MS further performed recruiting, conducted the study and the computations of the results. DP and CS provided input and critical feedback concerning the intrinsic motivation methodology as well as analysis. All authors discussed the results and contributed to the final manuscript.

Funding

MS and DP acknowledge support by the H2020-641321 socSMCs FET Proactive project. CS is funded by the EU Horizon 2020 program / Marie Sklodowska-Curie grant 705643.

Acknowledgments

We would like to thank Rebecca Miko for her help with recruiting participants and with her help on the final manuscript.

References

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016, November). Facets of the Fundamental Content Dimensions: Agency with Competence and Assertiveness—Communion with Warmth and Morality. *Frontiers in Psychology*, *7*. doi: 10.3389/fpsyg.2016.01810

Ay, N., Bernigau, H., Der, R., & Prokopenko, M. (2012, November). Information-driven self-organization: the dynamical system approach to autonomous robot behavior. *Theory in Biosciences*, *131*(3), 161–179. doi: 10.1007/s12064-011-0137-9

Ay, N., Bertschinger, N., Der, R., Güttler, F., & Olbrich, E. (2008, June 01). Predictive information and explorative behavior of autonomous robots. *The European Physical Journal B*, *63*(3), 329–339. doi: 10.1140/epjb/e2008-00175-0

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009, January). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics*, *1*(1), 71–81. doi: 10.1007/s12369-008-0001-3

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018, August). Social robots for education: A review. *Science Robotics*, *3*(21). doi: 10.1126/scirobotics.aat5954

Bialek, W., Nemenman, I., & Tishby, N. (2001, November). Predictability, Complexity, and Learning. *Neural Computation*, *13*(11), 2409–2463. doi: 10.1162/089976601753195969

Boden, M. A. (2008, February). Autonomy: What is it? *Biosystems*, *91*(2), 305–308. doi: 10.1016/j.biosystems.2007.07.003

Carpinella, C. M., Wyman, A. B., Perez, M. A., & Stroessner, S. J. (2017). The Robotic Social Attributes Scale (RoSAS): Development and Validation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 254–262). New York, NY, USA: ACM. doi: 10.1145/2909824.3020208

Castro-González, Á., Admoni, H., & Scassellati, B. (2016, June). Effects of form and motion on judgments of social robots' animacy, likability, trustworthiness and unpleasantness. *International Journal of Human-Computer Studies*, *90*, 27–38. doi: 10.1016/j.ijhcs.2016.02.004

Ceha, J., Chhibber, N., Goh, J., McDonald, C., Oudeyer, P.-Y., Kulić, D., & Law, E. (2019). Expression of Curiosity in Social Robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI'19.* ACM Press. doi: 10.1145/3290605.3300636

Christensen, H. I., Okamura, A., Mataric, M., Kumar, V., Hager, G., & Choset, H. (2016). *Next Generation Robotics* (Tech. Rep.). Retrieved from `https://arxiv.org/abs/1606.09205v1`

Clabaugh, C., & Matarić, M. (2019, May). Escaping Oz: Autonomy in Socially Assistive Robotics. *Annual Review of Control, Robotics, and Autonomous Systems*, *2*(1), 33–61. doi: 10.1146/annurev-control-060117-104911

Cohen, J. (1992, 7). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cover, T. M., & Thomas, J. A. (2012). *Elements of Information Theory.* John Wiley & Sons.

Crutchfield, J. P., & Young, K. (1989, July). Inferring statistical complexity. *Physical Review Letters*, *63*(2), 105–108. doi: 10.1103/physrevlett.63.105

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631–648. doi: 10.1037/0022-3514.92.4.631

Dautenhahn, K. (2004). Robots we like to live with?! - a developmental perspective on a personalized, life-long robot companion. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No.04TH8759).* IEEE. doi: 10.1109/roman.2004.1374720

Der, R., Güttler, F., & Ay, N. (2008, January). Predictive information and emergent cooperativity in a chain of mobile robots. In S. Bullock, J. Noble, R. Watson, & M. A. Bedau (Eds.), *The 11th International Conference on the Synthesis and Simulation of Living Systems (Artificial Life XI)* (pp. 166–172). MIT Press. Retrieved from `https://mitpress-request.mit.edu/sites/default/files/titles/alife/0262287196chap22.pdf`

Der, R., & Martius, G. (2006). From Motor Babbling to Purposive Actions: Emerging Self-exploration in a Dynamical Systems Approach to Early Robot Development. In S. Nolfi et al. (Eds.), *From Animals to Animats 9* (Vol. 4095, pp. 406–421). Springer Berlin Heidelberg. doi: 10.1007/11840541˙34

Der, R., & Martius, G. (2012). *The Playful Machine* (Vol. 15). Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-642-20253-7

Fiske, S. T. (2018, February). Stereotype Content: Warmth and Competence Endure. *Current Directions in Psychological Science*, *27*(2), 67–73. doi: 10.1177/0963721417738825

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007, feb). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. doi: 10.1016/j.tics.2006.11.005

Friston, K. (2010, January). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138. doi: 10.1038/nrn2787

Froese, T., & Ziemke, T. (2009, March). Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, *173*(3-4), 466–500. doi: 10.1016/j.artint.2008.12.001

Fukuda, H., & Ueda, K. (2010). Interaction with a Moving Object Affects One's Perception of Its Animacy. *International Journal of Social Robotics*, *2*(2), 187–193. doi: 10.1007/s12369-010-0045-z

Geukes, K., Breil, S. M., Hutteman, R., Nestler, S., Kfner, A. C. P., & Back, M. D. (2019, January). Explaining the longitudinal interplay of personality and social relationships in the laboratory and in the field: The PILS and the CONNECT study. *PLOS ONE*, *14*(1). doi: 10.1371/journal.pone.0210424

Gordon, G., Breazeal, C., & Engel, S. (2015). Can Children Catch Curiosity from a Social Robot? In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI'15.* ACM Press. doi: 10.1145/2696454.2696469

Grassberger, P. (1986, September). Toward a quantitative theory of self-generated complexity. *International Journal of Theoretical Physics*, *25*(9), 907–938. doi: 10.1007/bf00668821

Guckelsberger, C., Salge, C., & Colton, S. (2016, September). Intrinsically motivated general companion NPCs via Coupled Empowerment Maximisation. In *2016 IEEE Conference on Computational Intelligence and Games (CIG)* (pp. 150–157). IEEE. doi: 10.1109/cig.2016.7860406

Guckelsberger, C., Salge, C., & Togelius, J. (2018, August). New And Surprising Ways to Be Mean. In *2018 IEEE Conference on Computational Intelligence and Games (CIG)* (pp. 1–8). IEEE. doi: 10.1109/cig.2018.8490453

Hayashi, K., Shiomi, M., Kanda, T., & Hagita, N. (2010, September). Who is appropriate? A robot, human and mascot perform three troublesome tasks. In *19th International Symposium in Robot and Human Interactive Communication.* IEEE. doi: 10.1109/roman.2010.5598661

Huang, H.-M., Messina, E., Wade, R., English, R., Novak, B., & Albus, J. (2004). Autonomy Measures for Robots. In *ASME 2004 International Mechanical Engineering Congress and Exposition* (pp. 1241–1247). doi: 10.1115/IMECE2004-61812

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899–913. doi: 10.1037/0022-3514.89.6.899

Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., & Hagita, N. (2010, October). A Communication Robot in a Shopping Mall. *IEEE Transactions on Robotics*, *26*(5), 897–913. doi: 10.1109/tro.2010.2062550

Kaplan, F., & Oudeyer, P.-Y. (2004). Maximizing Learning Progress: An Internal Reward System for Development. In F. Iida, R. Pfeifer, L. Steels, & Y. Kuniyoshi (Eds.), *Embodied Artificial Intelligence* (Vol. 3139, pp. 259–270). Springer Berlin Heidelberg. doi: 10.1007/978-3-540-27833-7˙19

Klyubin, A. S., Polani, D., & Nehaniv, C. L. (2005). Empowerment: A Universal Agent-Centric Measure of Control. In *2005 IEEE Congress on Evolutionary Computation* (Vol. 1, pp. 128–135). IEEE. doi: 10.1109/cec.2005.1554676

Kulms, P., & Kopp, S. (2018, June). A Social Cognition Perspective on Human–Computer Trust: The Effect of Perceived Warmth and Competence on Trust in Decision-Making With Computers. *Frontiers in Digital Humanities*, *5*. doi: 10.3389/fdigh.2018.00014

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

Martius, G., Der, R., & Ay, N. (2013a). *Appendix with derivations and technical detail.* doi: https://doi.org/10.1371/journal.pone.0063400.s001

Martius, G., Der, R., & Ay, N. (2013b, May). Information Driven Self-Organization of Complex Robotic Behaviors. *PLOS One*, *8*(5), 1–14. doi: 10.1371/journal.pone.0063400

Martius, G., Jahn, L., Hauser, H., & Hafner, V. V. (2014). Self-exploration of the Stumpy Robot with Predictive Information Maximization. In *From Animals to Animats 13* (pp. 32–42). Springer International Publishing. doi: 10.1007/978-3-319-08864-8˙4

Maturana, H. R., & Varela, F. J. (1991). *Autopoiesis and Cognition: The Realization of the Living* (Vol. 42; R. S. Cohen & M. W. Wartofsky, Eds.). Springer Science & Business Media. doi: 10.1007/978-94-009-8947-4

Merrick, K. E., & Maher, M. L. (2009). *Motivated Reinforcement Learning* (1st ed.). Springer Berlin Heidelberg. doi: 10.1007/978-3-540-89187-1

Mieczkowski, H., Liu, S. X., Hancock, J., & Reeves, B. (2019, March). Helping Not Hurting: Applying the Stereotype Content Model and BIAS Map to Social Robotics. In *2019 14th ACM/IEEE international conference on human-robot interaction (HRI).* IEEE. doi: 10.1109/hri.2019.8673307

O'Regan, J. K., & Noë, A. (2001, October). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, *24*(5), 939–973. doi: 10.1017/S0140525X01000115

Oudeyer, P.-Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning. In *Motivation - theory, neurobiology and applications* (pp. 257–284). Elsevier. doi: 10.1016/bs.pbr.2016.05.005

Oudeyer, P.-Y., & Kaplan, F. (2008). How can we define intrinsic motivation ? In *Proceedings of the 8th international conference on epigenetic robotics: Modeling cognitive development in robotic systems.* Brighton, United Kingdom. Retrieved from https://hal.inria.fr/inria-00420175

Oudeyer, P.-Y., & Kaplan, F. (2009). What is Intrinsic Motivation? A Typology of Computational Approaches. *Frontiers in Neurorobotics*, *1*, 6. doi: 10.3389/neuro.12.006.2007

Oudeyer, P.-Y., Kaplan, F., & Hafner, V. V. (2007, April). Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Transactions on Evolutionary Computation*, *11*(2), 265–286. doi: 10.1109/tevc.2006.890271

Paolo, E. A. D. (2004). *Unbinding Biological Autonomy: Francisco Varela's Contributions to Artificial Life* (Vol. 10) (No. 3). MIT Press - Journals. doi: 10.1162/1064546041255566

Pinillos, R., Marcos, S., Feliz, R., Zalama, E., & Gómez-García-Bermejo, J. (2016, May). Long-term assessment of a service robot in a hotel environment. *Robotics and Autonomous Systems*, *79*, 40–57. doi: 10.1016/j.robot.2016.01.014

Ryan, R. M., & Deci, E. L. (2000a, 1). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67. doi: 10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2000b, January). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. doi: 10.1037/0003-066X.55.1.68

Scheunemann, M. M. (2019). *Intrinsically Motivated Autonomy in Human-Robot Interaction: Human Perception of Predictive Information in Robots.* Online. Retrieved 2019-10-10, from https://mms.ai/TAROS2019-supplementary/

Scheunemann, M. M., Dautenhahn, K., Salem, M., & Robins, B. (2016, August). Utilizing Bluetooth Low Energy to recognize proximity, touch and humans. In *2016 25th ieee international symposium on robot and human interactive communication (ro-man)* (pp. 362–367). IEEE. doi: 10.1109/ROMAN.2016.7745156

Scheunemann, M. M., Salge, C., & Dautenhahn, K. (2019). Intrinsically Motivated Autonomy in Human-Robot Interaction: Human Perception of Predictive Information in Robots. In K. Althoefer, J. Konstantinova, & K. Zhang (Eds.), *Towards autonomous robotic systems* (pp. 325–337). Cham: Springer International Publishing. doi: 10.1007/978-3-030-23807-0˙27

Schmidhuber, J. (1991). Curious model-building control systems. In *[proceedings] 1991 ieee international joint conference on neural networks* (pp. 1458–1463). IEEE. doi: 10.1109/ijcnn.1991.170605

Smithers, T. (1997). Autonomy in Robots and Other Agents. *Brain and Cognition*, *34*(1), 88–106. doi: https://doi.org/10.1006/brcg.1997.0908

Steels, L. (2004). The Autotelic Principle. In *Embodied Artificial Intelligence* (Vol. 3139, pp. 231–242). Springer Berlin Heidelberg. doi: 10.1007/978-3-540-27833-7˙17

Stubbs, K., Hinds, P. J., & Wettergreen, D. (2007, March). Autonomy and Common Ground in Human-Robot Interaction: A Field Study. *IEEE Intelligent Systems*, *22*(2), 42-50. doi: 10.1109/MIS.2007.21

Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic Robotics.* MIT Press.

Tremoulet, P. D., & Feldman, J. (2000, aug). Perception of Animacy from the Motion of a Single Object. *Perception*, *29*(8), 943–951. doi: 10.1068/p3101

Van Rensburg, E. J. J., Rechnitzer, A., Causo, M. S., & Whittington, S. G. (2001). Self-averaging sequences in the statistical mechanics of random copolymers. *Journal of Physics A: Mathematical and General*, *34*(33), 6381. Retrieved from http://stacks.iop.org/ja/34/6381

Zahedi, K., Martius, G., & Ay, N. (2013). Linear combination of one-step predictive information with an external reward in an episodic policy gradient setting: a critical analysis. *Frontiers in Psychology*, *4*. doi: 10.3389/fpsyg.2013.00801