

Avoiding Playfulness Gone Wrong: Exploring Multi-objective Reaching Motion Generation in a Social Robot

Martin Cooney¹ · Anita Sant'Anna¹

Accepted: 29 April 2017 / Published online: 26 May 2017
© The Author(s) 2017. This article is an open access publication

Abstract Companion robots will be able to perform useful tasks in homes and public places, while also providing entertainment through playful interactions. “Playful” here means fun, happy, and humorous. A challenge is that generating playful motions requires a non-trivial understanding of how people attribute meaning and intentions. The literature suggests that playfulness can lead to some undesired impressions such as that a robot is obnoxious, untrustworthy, unsafe, moving in a meaningless fashion, or boring. To generate playfulness while avoiding such typical failures, we proposed a model for the scenario of a robot arm reaching for an object: some simplified movement patterns such as sinusoids are structured toward appearing helpful, clear about goals, safe, and combining a degree of structure and anomaly. We integrated our model into a mathematical framework (CHOMP) and built a new robot, Kakapo, to perform dynamically generated motions. The results of an exploratory user experiment were positive, suggesting that: Our proposed system was perceived as playful over the course of several minutes. Also a better impression resulted compared with an alternative playful system which did not use our proposed heuristics; furthermore a negative effect was observed for several minutes after showing the alternative motions, suggesting that failures are important to avoid. And, an inverted u-shaped correlation was observed between motion length and degree of perceived playfulness, suggesting that motions should neither be too short or too long and that length is also a factor which can be considered when generating playful motions. A short follow-up study provided some additional support for the idea that playful motions which seek to avoid failures can

be perceived positively. Our intent is that these exploratory results will provide some insight for designing various playful robot motions, toward achieving some good interactions.

Keywords Entertainment robotics · Motion generation · Social robotics · Playfulness · Reaching

1 Introduction

1.1 Avoiding Playfulness Gone Wrong in Motions

This article reports on an exploratory effort to investigate how to generate good playful reaching motions in a home robot. Home robots, in addition to performing useful tasks like health monitoring and patrolling [1] or housework [2], can also seek to entertain by interacting in a playful way [3,4].

Playfulness is a typical trait [5] in humans and animals characterized by *fun*, *happiness*, and *humor* [6]: here fun means that a behavior is entertaining, enjoyable, or rewarding to an observer [7]. Happy means that a behavior indicates the performer is experiencing a positive state like subjective well-being [8]. And, humor means that a behavior comprises some anomalous aspect [9] within some recognizable framework [10,11] which seems “okay, acceptable or safe” [12], thereby affording anticipation [13] and insight [14]; this juxtaposition of recognizable structure and anomaly has been referred to as “method in madness” [15]. Fun, happiness, and humor are related but not identical. For example, watching a sad movie can be fun; happy behavior on a sad occasion might not be fun. A serious exercise program can be fun; a humorous blooper seen many times might not be fun. A child skipping can seem happy without humor, and slipping on a banana

✉ Martin Cooney
martin.daniel.cooney@gmail.com

¹ School of Information Technology, Halmstad University, Halmstad, Halland, Sweden

peel can seem humorous even if the person who slipped is not happy.

Playfulness in an appropriate context can have positive effects: fun is an important part of everyday life, even at work, and facilitates acceptance of technologies [16–19]. Happiness can elicit positive feelings in others through a process of emotional contagion [20]. Also, humor can positively affect well-being [21]; suggest empathy, trust [22], and intelligence [23]; and reduce social distance by indicating shared expectations and friendly intentions [24]. Positive effects have been observed additionally in Human–Robot Interaction (HRI): for example, in increased intention to interact with a fun robot [25], higher positive affect due to skillfully operating a happy-looking robot [26], good overall impression of two robots performing a humorous routine [27], and softened mood of interactions after a robot told jokes [28].

A robot can express playfulness in various ways, including through its motions. Embodied motion is an important modality to explore for robots because the ability to move in our world is a distinctive quality of robots compared to other technologies such as virtual agents, and it is particularly important for social robots intended to interact with people [29]. To design playful motions, some insight can be drawn from the literature. Fun-seeking is *autotelic* (conducted because it is intrinsically rewarding), so the fun “component” of a motion will not be functional or directed toward fulfilling some goal [30]. Another insight is that happiness has been reported to have been conveyed via curved motions [31]. (Why this might have been the case is not clear; we think that sharp motions which change direction suddenly could be unpredictable, which can be a feature of adversarial behavior in games, and that sharp edges in general might be associated with aggression, like in the seminal study by Heider and Simmel in which some moving triangles were seen as aggressive and a circle as playful and meek [32]). Also, humorous motions can involve moving in a way which has one unexpected aspect; for example, placing a banana somewhere is not humorous in itself but can become humorous if it would be placed on someone’s head [33]. Some similar effects were also observed in our previous work, in which translational motions in a robot were mostly perceived as goal-driven, whereas curved and repeated motions in a space without objects were perceived to be playful, and likened to tail-wagging in a pet [34,35].

A challenge is that attempts of a robot to appear playful can produce undesired results; for example:

- (Obnoxious) Some playful behavior was described as obnoxious. In one study, a robot asked participants to perform some exercises in either a serious or a playful manner, in the latter case joking with comments such as “Do you know how to breathe?” and laughing [36]. In another study, a robot was programmed to laugh during

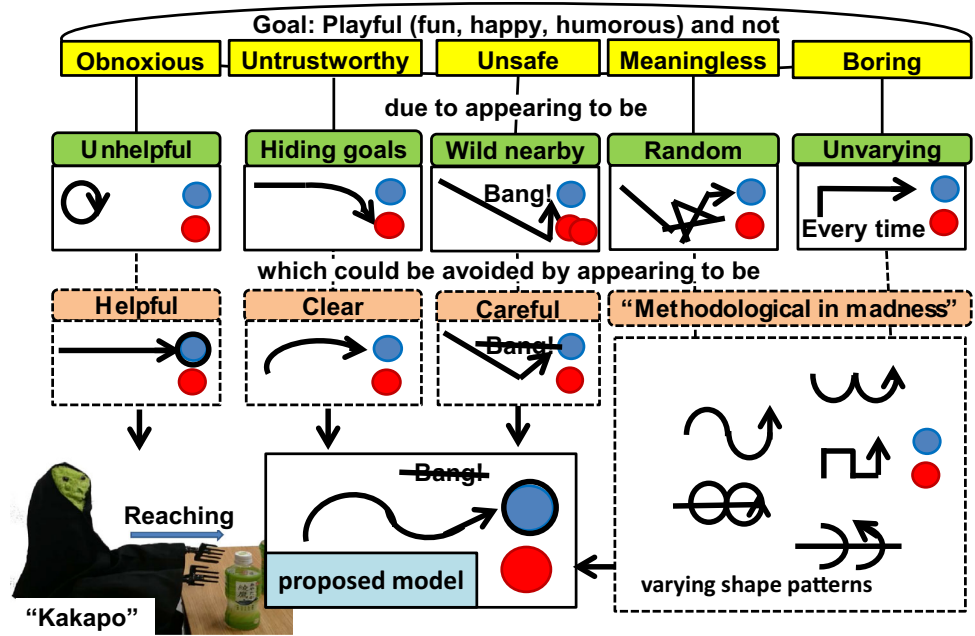
a game when an unfair offer was made to a participant, resulting in some negative impressions [37].

- (Untrustworthy) Playful motions which trick people have also elicited distrust. In one study, participants were asked to play a game by anticipating which object a robot arm was reaching for and pull it away; when tricked by the robot arm suddenly changing its trajectory, some participants felt less trust to the robot [38]. Another study hypothesized that consequences of such behavior could be serious: “once a robot cheats against a human to its advantage, it is difficult, even impossible, to forget” [39].
- (Unsafe) Playful interactions can also occur at close distances and involve spontaneity, but close-by unexpected motions can frighten. In one study, a child was afraid when a quadcopter launched itself from a racquet which the child was holding, or when it crashed or moved in an unexpected fashion [40].
- (Meaningless or Boring) In our own previous work, some playful motions seemed meaningless, which led to decreased interest in interacting [41]. Also, we used scripted motions which were played when a small robot was raised or lowered, laid down to sleep, made to walk or dance, hugged, or turned upside down, but over time playful behavior observed repeatedly can become boring. To avoid this problem, in one study a robot at an elementary school was programmed to gradually show more behavior over time, but a child who had missed the robot’s earlier progression found it boring to play with [42].

The question becomes: how can robotic motions be generated which appear *playful* but not *obnoxious*, *untrustworthy*, *unsafe*, *meaningless*, or *boring*? Some previous HRI studies have sought to identify typical failures and heuristics to avoid or mitigate them. For example, Satake et al. proposed how a mobile robot could approach humans successfully, reporting failures in which a robot’s behavior was not sufficient to attract attention, the robot did not respond in time, or the robot approached people who were too busy [43]. Lee et al. found that warning about a robot’s limitations beforehand and apologizing afterwards allowed mitigation of failures for a service robot delivering a wrong drink [44]. Also, our own previous work, for the context of playing with a small held robot via touch, proposed some heuristics for manually designed motions such as incorporating large motions to create the impression that a human’s interaction greatly affects the robot, and repeated suggestions to convey that the robot has a persisting intention [41]. However, the literature did not indicate how robotic motions can be generated which appear playful but not obnoxious, untrustworthy, unsafe, meaningless, and boring.

We derived ideas regarding possible causes for the reported failures from our experience and the literature: For

Fig. 1 Basic focus of the current work: exploring a model for generating playful reaching motions which avoids typical failures, using our new robot, Kakapo (*lower left*), and a motion generation framework, CHOMP: we proposed that designers can seek to generate motions which are helpful in completing a task, clear about goals, careful near objects and people, and featuring anomaly within a recognizable framework



obnoxiousness we believe that laughing at a human could be interpreted as derisive [45] and not expressing a genuine desire to support or help. Distrust could result when deceptive motions appear to have hidden goals. Safety of interacting persons is a priority and current robots are not always capable of performing flawlessly, which could lead to doubts, especially during a first interaction, if a robot moves wildly near a human or object. It might not have been conveyed clearly enough in our previous work that the robot was not just acting randomly but like a human or animal had some underlying consistency or “method” underlying its behavior. Scripts comprising a limited number of motions will not be feasible for long-term interactions because repetition and unreactivity evoke boredom [46]. Thus, failures could result from over-playfulness, hidden goals, wildness, randomness and lack of variety.

Based on our ideas, some good playful motions could be generated by conveying that a robot is helping the human by completing tasks, that its goals are clear and it has no other hidden goals, that it can move carefully near people and objects, and that the motion possesses some mixture of method and madness (structure and anomaly) which can be dynamically generated. Figure 1 summarizes this basic concept for the current work.

What was unclear was if our ideas would allow us to generate playful motions which avoid failures, or even if we could reproduce reported failures, due to the complexity of the phenomenon of perceived playfulness (it is not obvious even for humans how to generate playful motions), as well as restricted expressive capability in robots and potential differences in attitudes toward robots compared to humans and animals. For example, it is not clear that a robot could be

perceived as too playful because entertainment robots have been built which only exhibit playfulness. Likewise, humor is highly subjective and could be perceived even in a serious motion intended to deceive. Moreover, it might seem intuitive that a robot should never appear to move wildly, but there is a need in humans for outlets of chaotic play which can be messy and dangerous [47], thereby eliciting positive emotions by excitation-transference [48] and catharsis by the resolution of generated conflicts [49]. Random behavior exhibiting some similarity in form would also seem to satisfy some criteria for playfulness in living beings [50]. And, it could also appear obvious that a robot should not repeat itself, but some degree of repetition might be desired to “set up” anomalies and to clearly convey the robot’s intentions.

Thus, based on our ideas, we wished to investigate how people perceive motions, of which many different kinds are possible. We decided to focus our investigation on one kind of motion which is typical for assistive settings—reaching—to determine if it could be made to appear playful while avoiding failures.

1.2 Playful Reaching

To gain insight into how reaching motions could be incorporated into a playful social interaction, we came up with a narrative, below, and considered some additional properties of shape and length.

Max was not happy when the new robot Kakapo arrived at the care center: they said it would be able to do various tasks like pouring them juice, but what if he made some mistake or didn’t respond quickly enough and broke it? That first day, sure enough, someone had placed the apple juice, his apple

juice, out of reach next to Kakapo. “Say, if it’s not too much to ask could you get that for me?” he asked. The robot stopped wiping the table and its closest end effector rose to point at him then dipped down and back up in what he guessed was a large nod. It flew up majestically, twirling and looping and settled on the grape juice. “Okay, they must have used purple apples for that one,” Max said, smiling, “I think I’ll just have some of the regular apple juice over there, thank you.” Kakapo appeared to shrug, and again with a great flourish, it grasped the apple juice, then suddenly buckled down under the weight. “Are you okay?” Max asked alarmed. But it was only a show; the robot recovered immediately, bringing the juice to him in a way that reminded him of a rabbit bouncing. Max laughed, relieved. “Okay, you got me today,” he said, “but I’ve got my eye on you.” Kakapo nodded again, and he thought he saw some mischievous glint in its end effector.

The narrative indicated that our basic desired scenario, dyadic with one robot and one human, will not consist only of playful reaching. For example, Kakapo will also require the ability to recognize various signals and infer what people know (e.g., pretending that an object is heavy will not be perceived as playful if the person does not know it is light); to conduct simple communicative cues such as nods and shrugs as well as task-related motions such as grasping and carrying; and to be playful in relation to objects (e.g., taking into consideration size, weight, proximity, phase, color, hardness, location, temperature, or stiffness). Thus, such an interaction will require many components to come together to make sense, within an appropriate context. We decided to focus on generating playful reaching motions as our starting point because this was the fundamental/most challenging part for us which had not yet been addressed in the literature (recognition is very important but it is not sufficient to set up a reaching interaction). And, we assumed that people will have some idea that a robot they are interacting with could be playful (e.g. people might see a description on a robot’s package when they purchase it).

Furthermore, we chose to focus on the basic case of trajectories in two dimensions (particularly their shape and length), and to consider each proposed heuristic of our model to determine how playful reaching motions can be generated. First, for a robot to be helpful, its motion should reach the goal object. The simplest such motion can be defined as a straight shortest translation from start to goal. Next, the motion should be clear about goals. If there are other objects near the goal object the trajectory can be structured to avoid reaching or pointing toward them. Third, the motion can be designed to avoid impressions that collisions might occur. Fourth, the motion can be designed to have some kind of recognizable anomaly, which can be varied over the course of an interaction in terms of shape and length.

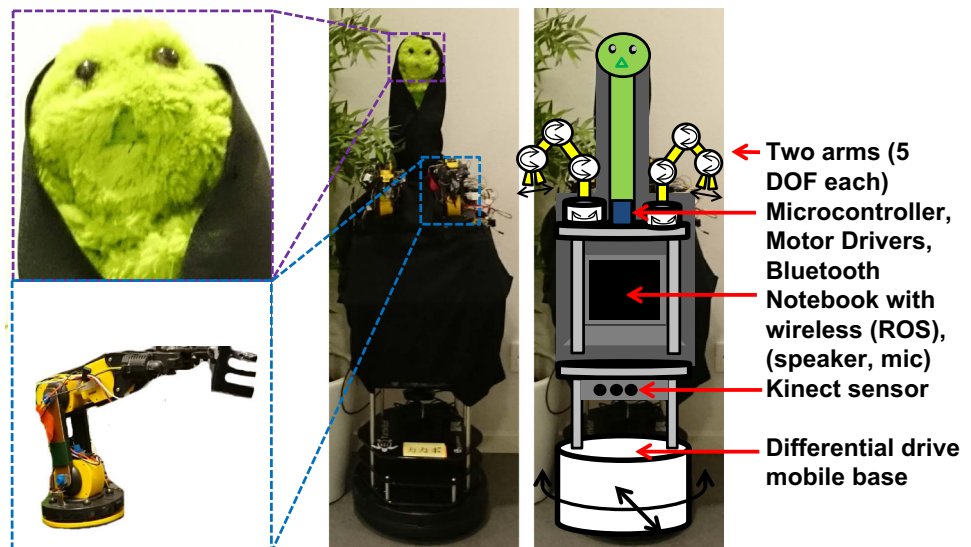
In the current work, for shape we selected some basic trajectories which could be easily generated mathematically and can also feature some curvature for conveying happiness: sinusoids, circles, trochoids, fast-transitioning waves and piecewise curves. Sinusoids are easily altered waveforms which have been much studied in engineering and mathematics. Circles may be attractive due to high symmetry [51]. Trochoids such as cycloids can be used to direct anomaly to one side of the axis from start to goal to reduce possible confusion about what the robot is reaching for. Waveforms with precise fast transitions such as square waves could be expected of a robot. And rich complex motions can be modeled using piecewise curves with various linear and nonlinear segments.

Similar trajectories have been used in another HRI study which investigated how a robot’s imitative motions might affect a human teacher [52]. In this study, participants were asked to perform eight “everyday” motions described as follows: how an airplane flies to an airport, how a phone is hung up, how a dog walks to a bowl, how an elevator descends, how an airplane does a loop, how to clean a window with a sponge, how a frog jumps, and how a feather falls. The airplane looping motion resembles the circle trajectory in the current work. The motions relating to elevators and feathers could be described as examples of fast-transitioned motions like a square wave. The remaining motions can be described as sinusoids or trochoids with varying frequency and amplitude.

In addition to shape, we were also interested how motion length might affect playfulness. This related to our proposal that playful motions should combine anomaly and structure, where it was not clear what the balance should be. We considered that the simplest non-playful motion would be the shortest straight trajectory minimizing the robot’s time and work. Then playful motions such as described above, would require at least this length, in addition to some extra length to convey anomaly. This suggested an attractive notion that the longer the motion, the more anomaly, and hence the more playfulness, could be shown. For the motion patterns suggested, this could be achieved by varying amplitude and frequency.

In summary, some past work has proposed heuristics for avoiding failures in robot interactions (e.g., [41, 43, 44]), and other work has been done on dynamically generating motions of a robot arm to be legible or deceptive [38, 53], but none of these studies had investigated how to generate playful reaching motions. The contribution of the current work is exploring how people perceive playful reaching motions of various shapes and lengths intended to also avoid failures based on appearing helpful, clear, careful, and combining structure and anomaly.

Fig. 2 Introducing a new playful platform we built, Kakapo: combining the face of an eponymous previous prototype, a Turtlebot base, and two modified toy arms controlled via Bluetooth



1.3 Kakapo and CHOMP

The components to implement included: a robot, and software capable of generating reaching motions designed to appear helpful, clear, careful, and feature “method in madness” via simple shape patterns such as sinusoids, circles, trochoids, squarewaves and piecewise curves. To perform reaching motions we built a new character-like robot platform, Kakapo, as shown in Fig. 2, by combining three components: the face of an eponymous previous prototype [54], a Turtlebot mobile base from Clearpath Robotics, and two toy arms augmented with potentiometers, motor drivers, and Bluetooth. The face and body were designed to appear playful, and child-sized in its relatively large eyes and short body (~1.25 m high), although this part of the robot’s appearance was not shown in the user study in the current work. The Turtlebot base, in conjunction with Robot Operating System (ROS), Festival, and CMU Pocket Sphinx, allows Kakapo to perform basic tasks which occur in interactions related to navigation within a mapped environment and speech-based communication. The toy arms are safe in the event of a failure due to the low power, low-torque motors and light plastic parts, which is desirable for interacting with humans, especially elderly or children, in close shared domestic environments.

In addition to a physical robot, software enabling dynamic generation of playful motions was required. We incorporated our model into a pre-existing multi-objective motion generation framework called Covariant Hamiltonian Optimization for Motion Planning (CHOMP) [55], which allows multiple objectives to be mixed on the fly. Thus the “purely useful” task of reaching for an object can be conducted while also fulfilling requirements of playfulness by varying parameters.

In the CHOMP approach, a trajectory is determined by randomly generating initial waypoints and iteratively updating their positions by gradient descent toward maximizing some objectives, until a stopping criterion is reached. Trajectory is considered to be a function mapping time to a configuration; thus the optimization criterion to minimize to generate playful motions is an objective *functional* mapping trajectories to real numbers, as follows:

$$U_{\text{playful}} = F_{\text{smooth}} + \omega_0 F_{\text{clear}} + \omega_1 F_{\text{obstacle}} + \omega_2 F_{\text{shape}} \quad (1)$$

F_{smooth} is a standard term in CHOMP intended to generate a smooth curved motion from start to goal (to be helpful and avoid obnoxiousness from not completing a task, the last waypoint is fixed to occur at the goal). F_{clear} can be used toward ensuring the robot’s goal is clear and avoiding a deceptive impression. F_{obstacle} like F_{smooth} is also a standard term in CHOMP allowing a robot to be careful and avoid collisions. F_{shape} is intended to provide some structured anomaly by combining the trajectory with some basic shape pattern such as a sine wave.

F_{clear} and F_{shape} were defined similarly in a simplified manner:

$$F = \frac{1}{2} \int (x(t) - s(t))^2 \quad (2)$$

where $x(t)$ is the current trajectory and $s(t)$ is a curve which possesses some desired characteristic (an appealing shape or distance from a distractor object). To find the gradient of F , the integral is approximated using a discrete summation with $\Delta t = 1$. We expand out of the binomial term and summation, regroup terms into a quadratic form ($Ax^2 + Bx + C$), and calculate the derivative:

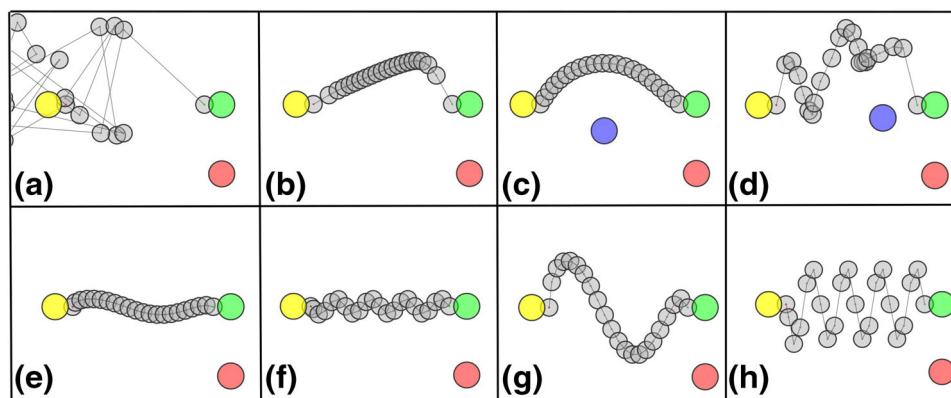


Fig. 3 Examples of reaching motions dynamically generated using playful CHOMP: **a** random initialization of waypoints (by default the gesture ends at the goal, ensuring motions are “helpful”), **b** showing clear goals by avoiding a distractor object, **c** showing also carefulness by avoiding an obstacle, **d** showing also anomaly using a sine wave pat-

tern (an example of the proposed system), **e–h** showing how sine wave trajectories can be dynamically varied to have high or low frequencies and amplitudes, which might allow expressing higher or lower degrees of anomaly/playfulness

$$\nabla F = (x(1)s(1); x(2)s(2); x(n-1)s(n-1)) \quad (3)$$

Figure 3 shows an example of some generated motions using a sinewave equation for $s(t)$ in F_{shape} , $s(t) = a \sin(ft)$ where a and f are parameters controlling amplitude and frequency. (A rotation matrix can be used to rotate the parameterized sinewave through an angle found via a dot product if the goal is not to the right of the starting point, as it is in our simplified scenario.) For $F_{autotely}$ the robot can be made to move more in regions where there is no distractor. We implemented this idea by fitting an approximating (B) spline through the start, a point found by reflecting the distractor location across the axis from start to goal, and the goal (two stubs were used for the remaining control points). The reflected point furthermore can be translated farther from the distractor via a parameter. Some example parameters which can be used are: number of waypoints $nq = 20$, weights $\omega_0 = 1.0$, and $\omega_1 = \omega_2 = 0.05$, convergence rate $eta = 100.0$, frequency $f = 1.0$, and amplitude $a = 2.0$.

1.4 Proposed Model and Hypotheses

In summary, we implemented a model intended to allow a robot to provide a good playful impression when reaching by lengthening the shortest trajectory to a goal via following some heuristics. The heuristics—appearing helpful, clear about goals, careful, and featuring structured anomaly—were described in Sect. 1.1, additional considerations of shape and length in Sect. 1.2, and our implementation in Sect. 1.3. Unclear was how the system would be perceived in regard to playfulness and avoiding failures, and how a motion’s length would relate to the degree of playfulness perceived. Three hypotheses were formulated:

- H1: (Playfulness) The proposed system will be perceived as playful.
- H2: (Failures) The proposed system will avoid some failures negatively impacting perceived goodness compared to an alternative playful system which does not take into account the proposed heuristics.
- H3: (Length) Motion length will exhibit a positive linear correlation with playfulness.

In H1–H3 above, the “proposed system” uses the different shape patterns described previously: sinusoids, circles, trochoids, square waves and piecewise curves. In H2, the failures referred to are appearing obnoxious, untrustworthy, unsafe, meaningless, or boring, which could come from over-playfulness, hidden goals, wildness, randomness, and lack of variety. Furthermore, for H1–H2 we were interested in both single motions and motion sequences. Also, H1 was formulated without comparison to another system because our goal was not to improve on the state of the art by making more playful motions than have been previously designed, but rather to imbue a functional reaching motion with properties that will allow it to appear playful while also satisfying some other expectations people have. H2 required testing because we did not have any evidence to suggest that the proposed system we had built would be non-failing, that our ideas for potential causes for reported failures might actually lead to failing, or even what the consequence of failing might be. For example, in the proposed system any deviation from a functional motion could have been perceived as irritating, or the simple fixed motion patterns like sinusoids could have been boring. In the other system, a motion which does not reach the goal could be interpreted as like a pet playing (e.g., chasing its own tail) which might not be considered failing.

And, regarding consequences of failing, if people just forgive a robot, it might not be necessary to avoid failures. We expected H3 would be fruitful because a simple relationship between motion length and playfulness would be convenient for designers to allow them to generate motions of various degrees of playfulness given a context.

2 Methods

2.1 Experiment Design

To verify our hypotheses from the previous section, we designed a user study, taking into account some additional requirements: We wanted to be able to gain a wide range of feedback, both qualitative and quantitative. We wanted to ease the task of evaluating robot motions for the participants to obtain fine feedback on various kinds of motions. We wanted to gain insight into how reaching motions are perceived in general, avoiding results associated only with the appearance of our particular platform. And, we desired some other system to compare with the proposed system which might be informative about failures.

To obtain various kinds of feedback we used the think-aloud method, questionnaires, and continuous evaluations. The think-aloud method was used because we did not know what kinds of qualitative impressions would result [56]. Questionnaires were used to obtain an overall quantitative evaluation of the proposed system and a different system used for comparison. A continuous evaluation method was used because we wanted to know how perceptions of playfulness would change over time and why, which required exploring the sequence as a process and not merely as a single event [57]. The think-aloud approach, questionnaires with similar items, and continuous evaluation methods have been used previously in HRI (e.g., [58–60]).

To facilitate evaluation, we adopted a “thin slice” approach of showing participants videos comprising short sequences, an approach which has been used previously in HRI [61]. It must be acknowledged that video clips do not allow for the same experience as watching an actual robot. However, videos possess some useful features as a prototyping approach for an exploratory HRI study: results in previous studies have been comparable with those from live trials, there is greater control over methodologies (exactly the same robot behavior can be shown, from the same perspective and lighting), they are quick to conduct and easy to change, and useful for avoiding wasting effort and checking assumptions [62]. In our case, we were interested in exploring how qualitative differences in stimuli are perceived, for which the usefulness of such short video segments has been described [63]. Moreover we wished to also explore perceptions of wild playfulness; our earlier research indicated that partic-

ipants can attribute danger to scenes in videos (depicting flying robots nearly colliding with a person’s head) and ethically we did not want participants to truly feel in danger, which could have occurred with a wildly moving robot in their close proximity. The videos also allowed us to employ a set up which was used previously for continuous scoring in HRI [60], in which participants scored motions while seeing a scoring grid and the most recent scores they had attributed. “Recent scores” refers to the scores given by one given participant for a motion sequence being viewed over several seconds. The scoring grid stated what the participants should evaluate, helping them not to forget, and had a column of numbers indicating each possible score (1–5). Showing the current and recent scores gave immediate feedback on what had been pressed and indicated that key presses had been recognized by the system. To do so without videos might have required a head mounted display which we did not have, and augmented reality could be perceived as unfamiliar or unnatural as this technology is not yet commonplace.

For the third requirement of selecting an experiment design, we felt that the important aim was to get feedback on an actual robot we had built, to be able to move forward and build some nice interactions, so we didn’t want to use simulations, and we wanted participants to know they were watching a robot. On the other hand, an experiment design that might favor the acquisition of generalizable results seemed useful for Kakapo with its distinctive appearance (e.g., its green face and yellow arms); therefore to reduce the effect of appearance we adopted an approach inspired by the “point light displays” used in human science. Johansson introduced this technique to be able to study how people perceive the visual information from motions while removing interference from extraneous factors such as form, color, and size; he reported that compelling impressions were evoked, yielding highly positive results [64]. In HRI a greater need for disentangling the effects of motion and form might exist because the space of possible forms is more expansive for robots than humans, suggesting that this approach should be explored. In such an approach, motions are represented by a few bright points moving in tandem. Using such a representation did not mean that the impression of a robotic arm was lost. We generated motions by videotaping Kakapo’s arm outfitted with light emitting diodes (LEDs) in a dark room. Thus the motions were actual robot motions, the arm can be seen faintly in the videos, participants were shown the arm before the experiment, and the instructions referred to the lights as representing the motion of the robotic arm. Therefore participants knew that they were watching a robot, but emphasis was placed on the motions and not on other factors such as colors and size of the embodiment.

For the fourth requirement for an informative system to compare to, our goal was not to show that we could

design playful motions for a robot (which we had explored in some previous work), but rather it was to verify that we could design reaching motions which would be perceived as both playful and good, avoiding failures. Therefore we also designed an alternative system which was intended to be playful but not necessarily good based on our ideas described in the Introduction section, although we also included some short straight motions in the videos to provide a basic indication that our system was working as expected.

2.2 Participants

The study was conducted with 14 participants at a university in Sweden (five females and nine males; average age = 30.5 years, SD = 10.2 years). Participants provided written consent stating that they wished to participate and that their data could be used. They had never interacted with our robot before and received no monetary compensation. Although the number of participants is not high, we felt that it was sufficient for starting to investigate impressions, given the exploratory nature of the current work.

2.3 Procedure

Participants watched and evaluated videos showing *point light display* sequences generated from playful reaching motions of a robot arm, over approximately one hour. Participants first entered a room and sat in front of a monitor. The door was closed, and only the participant and experimenter were in the room. Participants read some short instructions regarding the experiment scenario and their task. The videos were described as showing a robot arm in a home which would repeatedly reach from a start position to a goal position, a definition of playfulness was provided as “fun, happy, or humorous” with some examples, participants were told to disregard some slight changes in focus which resulted in lights appearing to become larger or smaller, and participants’ questions were answered. To reduce potential effects of novelty and anxiety, participants practiced watching and evaluating practice videos of the arm performing some simple motions. Then participants watched four videos. During the first two videos, participants verbally described each motion and completed a questionnaire after each video. During the latter two videos participants at any time freely pressed on numbered keys on a keyboard to score videos, yielding a continuous evaluation. Short interviews were conducted at the end.

2.4 Videos

Point light displays were generated in a darkened room by videotaping Kakapo’s right arm outfitted with LEDs of various colors: A yellow light on the end effector was used to

represent the robot arm. A green light represented an object which the robot arm was to grasp. A red light represented a different “distractor” object, used to test the effect of deceptive motions. And, a blue light was shown when the robot arm completed its task. Color choices were arbitrary and constant throughout the experiment. To generate motions, commands were transmitted via Bluetooth from a laptop to a microcontroller running code we wrote to interpolate and position the end effector by controlling four actuators. This set-up was not perfectly accurate, but we considered it to be sufficient to generate the kinds of motions we wished to show.

Using Kakapo’s arm, six videos were generated (two practice videos and four test videos). Videos will be made available online to facilitate reproducibility. A full HD camcorder (JVC GZ-HM40BU) with auto-focus and Konica Minolta HD lens ($f = 2.9\text{--}116\text{mm } 1:1.8$) was used to capture videos, which were shown in H.264/MPEG-4 AVC format. For each motion the end effector of the arm reached for a goal object, moving from its start position to an end position. Distances from start to goal for the end effector and from goal to distractor object were 12.5 and 6.5 cm, with an on-screen distance and speed of 7 cm and 5 cm/s. A top-down perspective from a camera directly above the arm was selected with consideration of the importance of orientation of point light displays [65]. Also, in videos V1 and V2, motions were separated by short segments of black video.

Videos V1–4 were intended to show changes in independent variables as follows:

- motion type:
 - proposed motions intended to be playful and avoid failures
 - non-playful straight motions
 - alternative motions intended to be playful but not take into account the proposed heuristics
- motion length (~1–20s)

Additionally we hoped to get some insight into how the shape patterns would be perceived, as well as how motions would be perceived when alone or in sequences. Thus, videos were structured as follows:

- V0: Practice videos: The robot moved to trace straight lines, triangles, and rectangles, to allow participants to practice thinking out loud and continuously evaluating.
- V1: Motions generated based on the proposed method:
 - V1.1 Sinusoid: The robot moved along an s-shaped trajectory to the goal.
 - V1.2 Circle: The robot moved along a looping trajectory to the goal.

- V1.3 Trochoid: The robot moved in short hops to the goal.
- V1.4 Square wave: The robot moved along a square wave trajectory to the goal.
- V1.5 Piecewise: The robot moved straight then horizontally, in alternating fashion, to the goal.
- V2: Alternative playful motions for comparison (which did not take into account the proposed heuristics for avoiding failures)
 - V2.1 Too-playful: The robot spun around once, showing only playfulness and not completing the reaching task.
 - V2.2 Hidden-goal: The robot moved first toward the goal, then switched at the end of its trajectory to reach for the distractor object.
 - V2.3 Rough-and-Tumble: The robot moved wildly, bumping into the goal object.
 - V2.4 Random: The robot moved randomly before reaching the goal.
 - V2.5 Repeated: The robot performed one scripted motion ten times without changes.
- V3: Motions varying in length
 - Motions V1.1.-1.5 were adapted to fill several minutes
- V4: A mixture of proposed and naive motions
 - Motions V1.1.-1.5 were used in the beginning and end, with naive motions in the middle.

For V3, to create a sequence capable of sustaining a perception of playfulness over a few minutes, we applied some additional heuristics: the rule of three, running gags, pauses, and emblems. The rule of three indicates that triplets are good for establishing short enjoyable patterns, and suggests this is why they are used in many contexts such as the typical three-act structure for script writing and various famous quotations such as “I came, I saw, I conquered” [66]. Running gags, based on the idea that a playful element can be repeated in several “callbacks” for good effect, relate to balancing method and madness. Pauses can offer “breathing room” between stimuli. An emblem is a gesture or motion which can represent a word, like waving a hand can express a greeting; termed “bisociation”, sudden recognition of a known pattern like an emblem in a new context can yield enjoyment and offers an explanation for why puns can be funny [67].

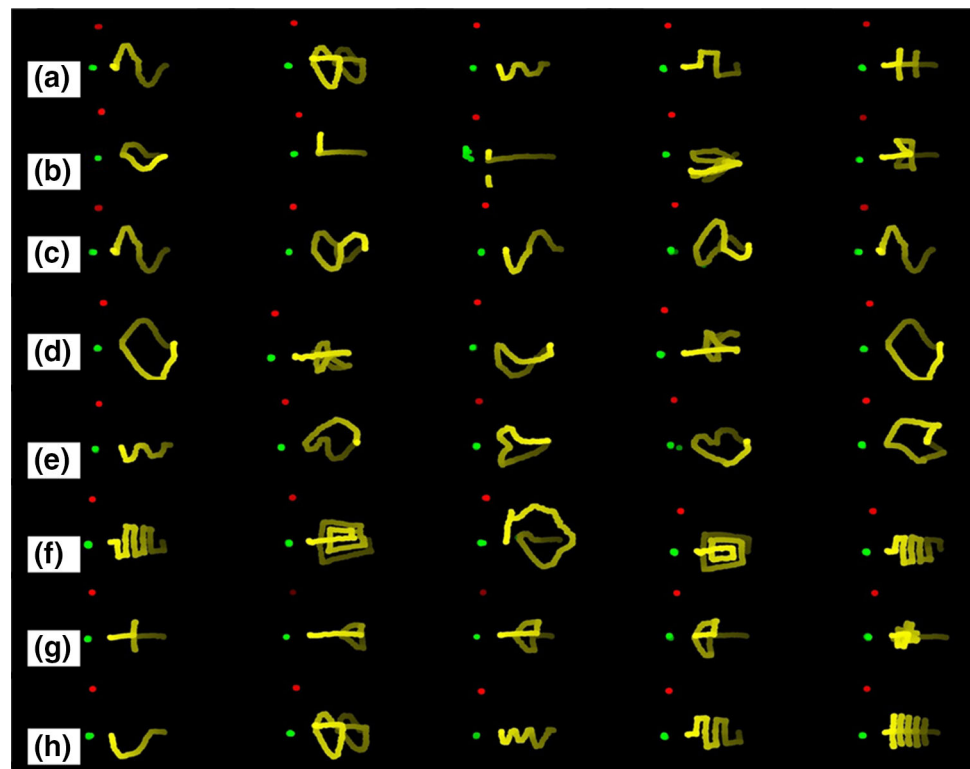
To structure the motion sequence, the video was broken down into an introduction, middle, and end, as in the rule of three, with the base motions in the beginning (sine, etc.) brought back at the end to provide a complete feeling. The base motions were also adapted as a form of running gag to

build five short episodes in the middle. To design the content of episodes inspiration was taken from fireworks and string figures, which involve play and shape. In some firework displays, a design like a heart is shown from different angles to facilitate observation from different perspectives; if a shape is not recognized at first it may be recognized the next time. In string figures, new shapes are progressively made from old shapes. Both art forms use recognizable shapes as emblems, in which the physical shape is associated with a symbolic meaning (e.g., a heart shape for love). To separate sequences, offer “breathing room”, and act as a simplified baseline, the straight shortest reaching motion was included at the start, middle, and end of the video, and between episodes.

Also, for further insight we varied some factors: the timing of “punchlines”, challenge of interpreting motions, and the type of sounds used to accompany motions. The positioning of motions of highest playfulness (punchlines, twists, or payloads) within each episode was varied. The sinusoid episode was designed to be playful throughout, the circle and trochoid episodes were most playful at the start, the edged episode had a twist in the middle and the piecewise episode peaked at the end. Also, challenge matched with skill can afford enjoyment [30]; therefore, the challenge required to understand what the robot was doing was varied to accommodate different levels of skill, making some episodes easier and some more challenging to understand. The sine and edged episodes were designed to be easy to understand, the piecewise episode was of medium difficulty, and the circle and trochoid episodes were harder. Some simple sounds were also used in V3 because we expect a robot to make sounds, although words were not used (the robot was restricted to utter “vwhee”). For the sine and piecewise episodes sounds were used which were contingent with the motions; for the circle episode melodic sounds were used, and for the trochoid and edged episodes repeated sounds were used. Sounds were not used in the other videos. Details were as follows:

- intro: all five basic shapes
- middle:
 - sinusoid: twisted back on itself to trace an infinity symbol (a lemniscate or figure-eight)
 - circle: shown as a sun-like symbol, followed by motions in the shape of a moon and star
 - trochoid: changed into a heart, which was rotated four times at 90°
 - square wave: changed into a spiral, a wacky spiral, a reversed spiral and a reversed square wave
 - piecewise: changed into a triangle rocket which moved closer to the goal then exploded
- final: all shapes shown (with length increasing toward forming a crescendo)

Fig. 4 Reaching motion trajectories: **a** proposed motions in V1, also used in V3 and V4, **b** failing motions in V2, **c–h** episodes and closing sequence for V3



V1–V2 were used to investigate questions H1–H2; V3 was used for H1 and H3; and V4 was used for H2. Speeds and timings of motions were similar and the videos edited after recording. Videos V1 and V2, and V3 and V4 were shown in random order. V1 and V2 were shown before V3 and V4 to avoid leading the participants and address our basic questions first. Trajectories, calculated using OpenCV to perform color detection from the videos, are shown in Fig. 4. V1.1–V1.5 are shown in Fig. 4a (each cell represents one reaching motion); V2.1–V2.5 are shown in Fig. 4b. V3 comprised all motions in the figure except (b). V4 comprised the motions in Fig. 4a at the beginning and end of the video, separated by a naive motion.

2.5 Measurements and Analysis

To evaluate the videos, the following measurements were made:

- M1: spoken thoughts (for think-aloud)
- M2: questionnaires
- M3: continuous evaluation scores
- M4: laughter
- M5: short interviews at the end

For the think-aloud task participants were asked to describe what the robot was doing and what kind of robot they thought it was. Questionnaire items were rated on a

scale of one to seven as in a Likert scale with (1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: neither, 5: somewhat agree, 6: agree, 7: strongly agree) in the form “The robot appeared”, for items Q1: likable, Q2: trustworthy, Q3: safe, Q4: purposeful, Q5: interesting, and Q6: playful. Likable was used as an antonym for obnoxious, trustworthy for untrustworthy, safe for unsafe, purposeful for meaningless, and interesting for boring. We used our own questionnaire because we could not find one which had the questions we wished to ask about playfulness and failing points, but the items have been used in previous studies (e.g., Q1–Q3 [59]). Positive items were used instead of five negative failure items and one positive item for uniformity. Participants could also freely note down impressions at the bottom of each questionnaire, which were used only to facilitate recall and were not analyzed. Continuous evaluations were also conducted, using a scale of one to five (with 1: strongly disagree, 2: somewhat disagree, 3: neutral, 4: somewhat agree, 5: strongly agree) for two statements: “This robot is currently being playful” (V3) “Overall this system seems good” (V4) The statement, scale, and evaluation values for a window of several seconds were shown below the videos as they played using OpenCV (we found that overlaying recent evaluation values could be distracting). Sounds were played via FMOD, a library for working with sounds (<http://www.fmod.org>). For additional insight into participants’ perceptions of playfulness, occurrences of laughter during video V3 were also counted. In the interviews at the end, participants were given a chance to

Fig. 5 Continuous playfulness score results for video V3 from all participants: the proposed system was perceived on average as playful for several minutes, compared to straight motions and a neutral score of 3; *dips* represent pauses in which straight motions were played

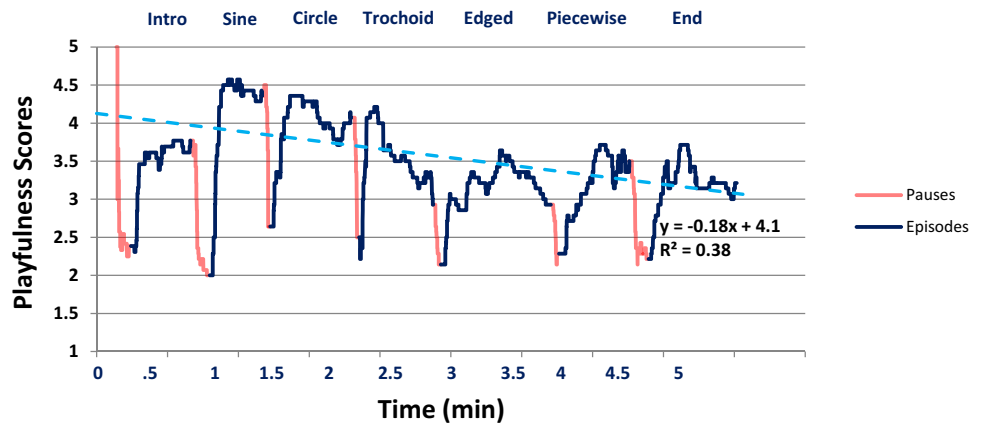
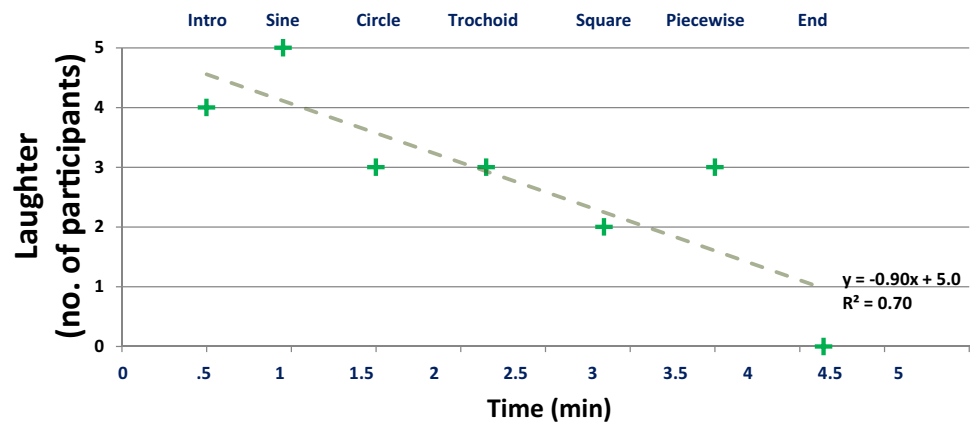


Fig. 6 Numbers of participants who laughed during each part of video V3 show a downward trend due to novelty



freely comment, and also asked if they had felt playfulness at any time during the experiment and why, and about the sounds.

Measurements M1–3 were analyzed using *t* tests, chi-squared tests, and graphs of average continuous evaluations over time. A statistical difference level of $\alpha = .02$ (.05/3) was selected as a Bonferroni correction for multiple hypotheses, because three hypotheses were investigated. Questionnaire items for failures were inverted and grouped to assess degree of failure. For M1, participants’ words were transcribed and coded by the experimenter. The numbers of positive and negative comments were counted, and typical comments made by more than one participant were also extracted. To check reliability one participant’s data were also coded by a second coder (Cohen’s Kappa = .74). Analysis on coded impressions was conducted via chi-squared tests. For M2, one-sample and paired *t* tests were used. For M3, one-sampled *t* tests, and paired *t* tests of average scores before and after failure occurred, were used. For M3 and M4, graphs were plotted, peaks identified and models fitted. For example, to gain a better understanding of the effect of length, the start and end points of motions in V3 (45 total, 10 straight, 35 proposed) were manually labelled and matched with playfulness scores. Comments in the interviews regarding playfulness and sounds were also coded.

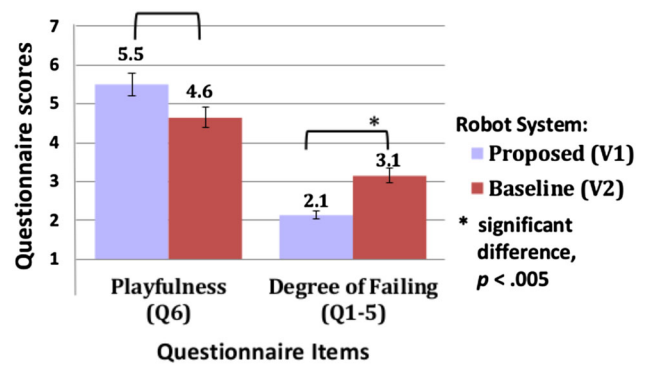


Fig. 7 Questionnaire results for videos V1 and V2: the proposed system was perceived as failing less than the alternative system

2.6 Results

Figures 5, 6, 7, 8, 9, and Table 1 show results relevant to each hypothesis.

H1 The proposed motions were perceived as more playful than straight motions (3.6 ± 0.43 vs. 2.3 ± 1.1 out of 5: $t(13) = 4.5, p = .001, d = 1.20$), with scores from both the questionnaires and continuous scoring also above neutral: (5.5 ± 1.2 vs. 4.0 and 3.6 ± 0.43 vs. 3.0): $t(13) = 4.8,$

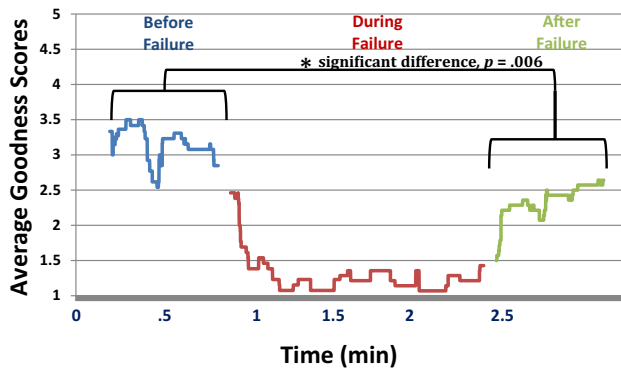


Fig. 8 Average goodness scores from all participants for continuous evaluations of video V4: after the naive motions, the robot was considered to be significantly worse than beforehand

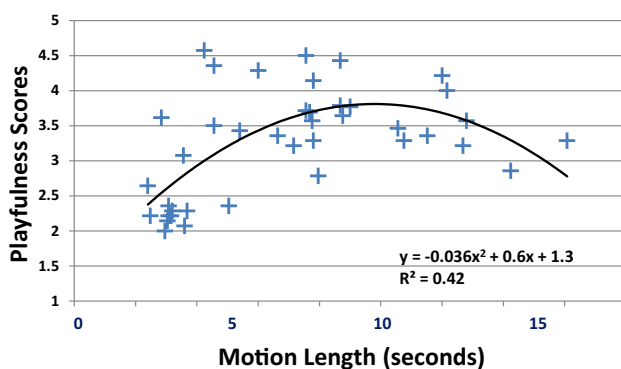


Fig. 9 Degree of perceived playfulness as a function of motion length with a second order polynomial trendline, based on evaluations of video V3

$p < .0005$, $d = 1.29$ ($<\alpha = .02$), $t(13) = 4.9$, $p < .0005$, $d = 1.30$. In addition 20 laughs were counted during the playful motion sequences of video V3 (1.7 ± 1.6 laughs per participant). Thus, the proposed motions, despite being functional in reaching for an object, were perceived as at least somewhat playful, in accordance with hypothesis H1, for both single motions and short sequences.

H2 The proposed system was perceived as failing significantly less than the alternative system both in terms of questionnaire scores: $t(69) = -5.8$, $p < .0005$, $d = 0.69$, and think-aloud comments according to a chi-squared test with Yates correction ($\chi^2(1, N = 35) = 19.3$, $p < .001$, $V = 0.74$). In the motion sequences mean goodness scores before and after alternative motions differed significantly (with one participant's data removed because he did not rate the robot at all before the alternative motions): $3.2 \pm .87$ before vs. 2.3 ± 1.2 after, $t(12) = 3.4$, $p = .006$, $d = 0.94$. Mean perceived goodness before the alternative motions (3.0 ± 1.0) was also higher than a score of 2.0 expressing slight disagreement that the robot performs well, $t(13) = 3.7$,

$p = .003$, $d = 1.32$, suggesting that playfulness was not regarded a failure in the system. (Scores at the start were conservative, possibly because playful motions are not the best possible motions if the criterion is reaching a goal quickly, and the participants didn't know what was going to happen later in the video.) In the interviews, nine participants said they would prefer the proposed system to the alternative (only one participant preferring the alternative system because he wanted to know why it was moving the way it did). Thus hypothesis H2, that the proposed motions would avoid some failures, was supported.

H3 Fig. 9 shows a graph comparing motion length to average playfulness score. Despite noise due to confounding factors (some motions were more interesting than others and novelty effects also were involved) a rough pattern appeared to exist. Fitting exponential, linear, logarithm, second-order polynomial, and power curves to data composed of motion lengths and average playfulness scores, the polynomial approximation provided a highest squared correlation value ($R = .42$ compared to linear $R = .19$, with $p < 0.01$ and a peak at $x = 256.25$ (8.6 s), $y = 3.9$). This interpretation also seemed to be supported by comments during the interviews: Participants indicated that straight motions appeared boring (3 participants, 21%) and not playful (21%). Short motions were good (50%), efficient (21%), safe (14%), goal-oriented (14%), and could incorporate some extra motion (14%); they could be playful (36%) if with sound (14%), or not playful (29%). Long motions could also be playful (14%) but were difficult to understand (29%), annoying (14%), boring (14%), and frustrating (14%). (Some ambiguity occurred due to participants' differing definitions about what constituted "short" or "long".) Thus results did not support hypothesis H3 predicting a simple linear increasing relationship, but suggested that motions intended to be both playful and good should be neither too short nor too long.

Additional Insight Further analysis was conducted to investigate how playfulness had been perceived. We compared the proposed and alternative systems, checked for patterns in continuous scores and laughter, and extracted typical impressions. The alternative motions in V2 were designed to be playful, without incorporating the heuristics we proposed for avoiding failures. To check, we conducted a chi-squared test with Yates correction on think-aloud comments, which confirmed that participants did not more frequently describe the proposed system as playful than the playful alternative system ($\chi^2(1, N = 17) = 1.0$, $p = .3$), and a statistical difference was also not observed in perceived playfulness scores on the questionnaire, 5.5 ± 1.2 vs. 4.6 ± 1.0 , $t(13) = 2.5$, $p = .03$ ($>\alpha = .02$). As shown in Fig. 5 and Fig. 6, average playfulness scores and numbers of participants who laughed during each part of video V3 decreased

Table 1 Typical impressions of motions obtained from think-aloud comments voiced by two or more participants (digits indicate the numbers of participants who voiced an impression)

Proposed			Alternative		
Motion	Comments	#	Motion	Comments	#
Sinusoids	+Good	3	Too-playful	–Not good (unclear)	3 (2)
	Natural	4		–Not playful	2
Circles	+Playful	2	Hidden-goal	+Playful	2
	Exploring	2		–Not good (problem, weird, unclear)	9 (5, 2, 2)
Trochoids	+Good	4	Rough-and-tumble	Natural	2
	+Playful	2		–Not good (problem)	4 (2)
	Deliberate	3			
Edged	+Good	2	Random	+Playful	2
	Mechanical	2		Exploring	6
	–Not good	2		–Not good (problem, unclear)	6 (5, 2)
Piecewise	+Good	2	Repeated	+Good	2
	+Playful	2		Deliberate	2
	Exploring	4		Exploring	2
	ECG-like	2		Mechanical	2
	Natural	2		–Not good (unclear, boring, inefficient)	8 (4, 2, 2)
	Mechanical	2			
	–Not playful	2		–Not playful	2

over the duration of five minutes; this may have been due to accustomization, a reduction in nervousness, or order effects.

The shape of the graph for Fig. 5 is intuitive in that peaks can be observed around the “punchlines”: the graph was flat during the sine phase which had punchlines throughout (the figure eight and last swoop), a decreasing slope emerged for the circle and trochoid episodes which had punchlines at the start (shapes and sounds were repeated), the peak for the edged episode occurs in the middle (when the spiral became large at once), and for the piecewise episode at the end (coinciding with the “explosion”). Both graphs also featured a maximum peak during the sine episode (composed of five participants, 25% of the total).

Think-aloud comments, shown in Table 1, suggested that some participants viewed curved motions as playful, whereas opinion was divided for the straighter piecewise motions. Interview comments offered additional insight, indicating the robot seemed playful when it moved forward then retracted (29%), showing a mix of caution and curiosity, or appeared animal-like (29%). Some felt the robot was playful most of the time (14%) or described incidents like an explosion (29%), zigzag (21%), or game similar to how parents move a spoon playfully to feed children (14%). Playfulness was difficult to perceive for some due to the abstraction of the point light display (14%), and during repetitive sequences (14%), which for one participant evoked an impression of

a person trying hard to be funny with jokes that are not clever.

Thus participants appear to have perceived the stimuli mostly as expected—participants attributed some playfulness to both the proposed and alternative playful systems in their words and scores, there was a reduction in playfulness attributed and laughter over time, and perception of playfulness coincided with punchlines—but there were also some unexpected perceptions such as mixed feelings about motions with straight components.

To gain further insight into perception of goodness, additional *t* tests were conducted on each failure questionnaire item, and perceived goodness after alternative motions was checked:

- Q1 Obnoxious: $1.9 \pm .92$ proposed vs. 2.7 ± 1.5 alternative, $t(13) = -2.2$, $p = .04$
- Q2 Untrustworthy: $1.9 \pm .62$ vs. 2.9 ± 1.3 , $t(13) = -3.2$, $p = .007$, $d = 0.85$
- Q3 Unsafe $2.5 \pm .94$ vs. 3.9 ± 1.6 , $t(13) = -3.3$, $p = .005$, $d = 0.89$
- Q4 Meaningless 2.4 ± 1.0 vs. 3.3 ± 2.1 , $t(13) = -2.0$, $p = .07$
- Q5 Boring: $1.7 \pm .61$ vs. 2.6 ± 1.2 , $t(13) = -2.3$, $p = .04$

Significant differences emerged for untrustworthy and unsafe, but not for obnoxious, meaningless, and boring. The rea-

son was the simplified experiment setting and short motion durations; some participants professed to liking the repetitive motion for a factory robot or thought it expressed a clear unchanging purpose and explained the robot's random motion as exploring. Think-aloud comments shown in Table 1 also suggest that participants saw both good and bad in the alternative motions. Mean perceived goodness after the alternative motions (2.3 ± 1.1) was not significantly lower than a neutral score of 3.0, $t(13) = -2.3$, $p = .04$, due to some mixed feelings (some participants thought it was good that the system could recover) and some unfamiliarity with continuous scoring and point light displays might have also played a role (some participants indicated that they had forgotten to consider everything they had seen and judged the goodness of the robot's current motions, and some described the abstractness of the LEDs and difficulty imagining that the robot had failed at a serious task). Thus, although the proposed system was perceived as failing less than the alternative, perceptions varied due to the complexity of the phenomenon.

On the side, we also investigated participants' abilities to perceive patterns in reaching motions and how sounds were perceived. No participants described seeing figure-eights, heavenly bodies (sun, moon, and stars) or hearts in V3. Also, for motion V2.5, repeated 10 times, eight participants described differences which did not exist (57%, total: 22 times, ave: 1.6 ± 2.0), in trajectory shape, timing, speed, distance, or the starting point for the motion. On average, participants indicated that motions were similar or the same after watching four times (ave: 4.0, 4.3), but only six participants at the end had voiced the conclusion that the motions were all the same (43%). Thus, participants were not able to recognize complex shapes in trajectories and pareidolia was noted in comments regarding perceived differences in identical motions, suggesting that designers should incorporate large differences in sequentially occurring motions to facilitate perception of anomaly. Sounds overall were described in the interviews as playful or good (according to 8 participants, 57% of total), especially when they changed in accordance with motions (50%), and exerted a strong effect (29%). However some sounds were not as good as others (50%), especially constant unchanging sounds (29%), and could be annoying (21%), weird or scary (14%). Participants liked the melodies (50%) and vehicle sounds (43%) but disliked buzzing (14%) and explosive sounds (14%). Thus, contingent and melodic sounds were perceived as playful and good.

Summary In summary, H1 and H2 were supported, with more positive impressions attributed to curved motions, but H3, predicting a positive linear correlation of playfulness with length, was not.

2.7 Follow-Up Study

The study described in the previous section was intended to provide a basic check on our proposed system and some exploratory insight, but there were various confounds. For example, the proposed system might have been rated as playful and good because participants knew that the study was about playful and good motions (there might have been good subject effects). Watching point light videos could be different from observing an actual robot. A within-subject design can suffer from carryover and practice effects which might not be fully prevented by counterbalancing (participants' ability to perceive the conditions can increase or decrease over time due to fatigue or learning). And, the think aloud method requires some mental effort (cognitive load) because participants must themselves come up with appropriate labels.

Therefore, to obtain some additional confirmation which tries to avoid these confounds, a small simplified follow-up study was conducted. In the follow-up study, participants were not told the purpose of the study, an actual robot and between-subject design were used, and the Geneva Emotion wheel (GEW) [68] was used to help participants to report their impressions. The study was conducted with 16 participants (5 females and 11 males; average age = 27.8 years, SD = 8.8 years) who had never interacted with a robot before (none of them were the same as in the previous study), and received no monetary compensation.

For simplicity in the follow-up study, a single hypothesis was formulated to summarize the core idea in the current article: a motion designed to be playful and good will evoke a more positive impression than motions which seek only to be playful or only to be good/functional. To test the hypothesis, each participant, over approximately five minutes, received instructions and gave consent, watched a motion and described their impression, and explained their impression in a short interview.

First, each participant was led into a room with only the participant and the experimenter, and the robot, which was situated in front of a table with two PET bottles (as in the previous study). There, participants were told only that they would watch a robot motion and give their impressions (e.g., they were not told that the study was about playful and good motions or that they should expect the robot to reach for a bottle). Then participants were asked to provide written consent stating that they wished to participate and that their data could be used. Next participants read a short handout on the GEW (the prescribed explanation from the authors' homepage at the University of Geneva at <http://affective-sciences.org/en/gew/>) and had an opportunity to ask questions about the GEW (no other questions, which might convey the purpose of the study or bias participants, were answered).

Table 2 Impressions of motions [format: total number of comments: emotion term (strength described by each participant)]

	Positive	Negative
C1 (playful only)	4: Amusement (4, 4), Joy (4), Pleasure (3)	4: Disgust (1), Fear (1), Disappointment (2), Boring
C2 (good only)	2: Interesting (4, 2)	3: Regret (3), Fear (2), Anger (1)
C3 (proposed)	10: Amusement (5, 3, 3, 2), Joy (5, 2, 4), Pleasure (3, 1), Interesting (2)	0

Then, participants stood in front of the robot, and the robot was triggered to reach for a bottle. The way the robot reached for the bottle was controlled as the independent variable of the experiment, and consisted of three conditions:

- C1: playful only: the robot arm moved in a sinusoid shape without reaching for a bottle.
- C2: good only: the robot arm moved directly in a straight line to reach for a bottle.
- C3: proposed (playful and good): the robot arm moved in a sinusoid shape to reach for a bottle.

Sinusoids were used for C1 and C3 with a simple sound (a “vwhee” with a short laugh) like in the previous study because this sequence was rated most playful and provoked most laughter in the previous study. A straight motion was used for C2 as a simplest example of a good reaching motion. The participants were divided roughly equally into three groups: six participants participated in C1, and five each in C2 and C3. After watching the motion, participants used the GEW to report their impressions. The GEW lists 20 impression categories (10 positive, 10 negative): Interest, Amusement, Pride, Joy, Pleasure, Contentment, Love, Admiration, Relief, Compassion, Sadness, Guilt, Regret, Shame, Disappointment, Fear, Disgust, Contempt, Hate, and Anger. Participants could select one or more impressions, or none, or use their own word to describe impressions if they felt there was no appropriate category.

Analysis was conducted by coding impressions as positive or negative and comparing between the number of comments made for each of the conditions. Two impressions were removed because they did not reflect the motions; the participants in the interviews said these impressions resulted from seeing a robot move for the first time and not from the motions themselves.

The acquired impressions are shown in Table 2. The interviews yielded some insight into some negative impressions reported for C1 and C2 but not for C3; C1 disappointed because it just played but didn’t try to do anything, whereas C2 seemed like a robot, in being cold and strict, and didn’t do anything in a special way. Thus, the basic idea of the current article, that positive impressions can be achieved by designing robot motions to be both playful and good, was again supported.

3 Discussion

3.1 Findings

The current work, extending our previous work on playful motions, explored how to generate some playful motions which are also perceived as good (not obnoxious, untrustworthy, unsafe, meaningless, or boring), in the context of reaching. The challenge was that human perception of playfulness is extremely complex; to derive a model and gain basic insight required bringing together insight from numerous sources in the human science and HRI literature and employing a mixture of methodologies (point light videos, think aloud, continuous scoring, and Geneva Emotion wheel). Technical improvements along the way also included extending an algorithm (CHOMP) to generate playful motions which avoid failures, and creating a new robot intended to appear playful and be able to safely engage in reaching interactions, Kakapo. The results of a conducted pilot study suggested that a playful impression could be sustained by motions and simple sounds for several minutes. Also a persisting negative effect on a robot’s perceived goodness was observed subsequent to conducting some motions which did not follow our model, suggesting that failures should be avoided. And, an inverse u-shaped curve was fit to data to model the relationship between motion length and playfulness with a sweet spot slightly before ten seconds, suggesting that motion length should also be considered when designing playful motions. A follow-up study seeking to avoid some confounds such as good subject effects yielded similar results, indicating that positive impressions can result from designing for both playfulness and goodness.

3.2 Possibilities for Combining Our Work with Others

We believe our results could be relevant for various home robots intended to perform useful tasks and companion robots. For example, our model could be used by a Roomba to playfully clean during a celebration, or a NAO to playfully provide an elderly person’s medication at the right time, or for dancing in a confined environment. The proposed model could possibly also be combined with expectancy-setting and recovery strategies to further mitigate adverse impressions due to failures [44]. Personalization could also

be interwoven into interactions by relating motions to a person or robot's behavior or preferences [69]; for example, in the context of reaching, a robot could hand a tired person a playful makeshift pillow, or shudder dramatically to show trepidation when reaching for an awkward object it might drop. Turn-taking is also likely to be important for longer interactions, so that a robot should act in such a way as to allow and entice humans to act [70]; for reaching, one playful interaction could emulate how Turkish ice cream vendors quickly offer and retract ice cream cones (<https://en.wikipedia.org/wiki/Dondurma>), or how cats sometimes seek to move their paws to be on top if something is placed over them. Another promising mechanism involves reinforcement [71]. For example, time spent interacting can be rewarded [42]. A robot could become gradually more affectionate toward a person, spending more time close by and reaching faster or more playfully, although such a strategy will require care to avoid turning away first-time users by showing only a fraction of the robot's repertoire, or punishing users who like the robot by only showing new motions to users who appear to lose interest. Also, human motion modeling approaches such as Laban Movement Analysis or kinesthetics may provide insight for how to generalize and model motions in higher dimensions.

3.3 Limitations

The results of the current exploratory work are limited by potential confounds arising from the participants in the user study (a small number of volunteers at a university), the target state to express (playfulness), the kind of motion (reaching), the mode of expression (videos of a single point of light on a robot's end effector moving in a simplified scenario), length of interaction (short), the categories of proposed and alternative motions (a few examples out of a vast space of possible motions), some Bokeh and persistence of vision effects, and ordering effects (in video sequences).

3.4 Future Work

More controlled psychological experiments with more participants from various countries and age groups will yield better understanding of key concepts. Furthermore, other beneficial expressive states will be investigated; for example, motions incorporating childlike qualities such as small rounded forms and a close proxemics distance could seem "cute", whereas rebelliousness, disengagement, roughness, and an aloof distance in regard to task, objects, and persons could be perceived as "cool". Moreover, other motions should be studied, such as those involving objects (e.g., how a robot could play with a string, ball, or rod), or locomotion, and on other robots; for example, clear communication of intentions will be useful for intelligent autonomous vehicles, but

additional requirements will be taken into account, such as not being able to move sideways directly or upwards or to be viewed from above.

Playfulness will also be investigated in artifacts other than robots, such as furniture or objects in a smart environment. For example, in a home, how could lamps, beds, chairs, or walls move playfully? A lamp could appear to become drowsy when its owner stays up late, and fall asleep when turned off. A bed could pretend to swallow a child, making munching sounds, and a chair could run from or chase a child. Shadows on a wall could dance a waltz with an elderly person when other company is absent.

Our next step, now that we have a system which can conduct playful reaching, will be to set up and investigate some interactions, which will also involve recognizing signals from people (are they being playful and is it okay if the robot is playful?) and leveraging the context (e.g., location and objects). We will also investigate how playfulness can be combined with other factors such as affection and conveyed in various modalities such as touch, sound/speech, and light. Our intent is that this knowledge will facilitate the design of playful robot motions which are "transparent" (clearly convey the robot's intentions) and thereby evoke desired impressions, toward achieving some good interactions.

Supporting Information

S1 Video

To clarify what was done A video showing, inter alia, some point light display motions used in the experiment will be made available on the internet.

S2 Code

To make tools available to others Open source code extending CHOMP for generating playful motions has been uploaded to the internet (<https://github.com/martincorney/trychomp>).

Acknowledgements A very large thank you to Roland Philippsen for help with the concept and CHOMP (math and code), to the volunteers who participated in the experiment, and to everyone else who helped. We received support from the Swedish Knowledge Foundation for the SIDUS AIR project.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Andreasson H, Magnusson M, Lilienthal AJ (2007) Has something changed here? Autonomous difference detection for security patrol robots. In: Proceedings of the IEEE/RSSJ international conference on intelligent robots and systems (IROS), pp 3429–3435
2. Forlizzi J, DiSalvo C (2006) Service robots in the domestic environment: a study of the Roomba vacuum in the home. In: Proceedings of the 1st ACM SIGCHI/SIGART conference on human–robot interaction (HRI), pp 258–265. doi:10.1145/1121241.1121286
3. Fujita M (2004) On activating human communications with pet-type robot AIBO. Proc IEEE 92(11):1804–1813. doi:10.1109/JPROC.2004.835364
4. Wada K, Shibata T (2007) Living with seal robots—its sociopsychological and physiological influences on the elderly at a care house. IEEE Trans Robot 23(5):972–980. doi:10.1109/TRO.2007.906261
5. Kuts E (2009) Playful user interfaces: literature review and model for analysis. Digital Games Research Association (DiGRA). Brunel University. Breaking New Ground: Innovation in Games, Play, Practice and Theory
6. playful (2010) Random House Kernerman Webster’s College Dictionary, Random House, Inc. [cited November 12, 2014]. <http://www.thefreedictionary.com/playful>
7. Winn B (2008) The design, play, and experience framework. In: Handbook of research on effective electronic gaming in education, vol 3, pp 1010–1024
8. Helliwell JF, Putnam RD (2004) The social context of well-being. Philos Trans R Soc Lond B Biol Sci 359:1435–1446. doi:10.1098/rstb.2004.1522
9. Webster J, Martocchio JJ (1992) Microcomputer playfulness: development of a measure with workplace implications. MIS Q 16(2):201–226
10. Valitutti A, Toivonen H, Gross O, Toivanen JM (2012) Decomposition and distribution of humorous effect in interactive systems. AAAI Technical Report FS-12-02 Artificial Intelligence of Humor
11. Fink EJ (2013) Writing the Simpsons: a case study of comic theory. J Film Video 65(1):43–55. doi:10.5406/jfilmvideo.65.1-2.0043
12. McGraw AP, Warren C, Williams L, Leonard B (2012) Too close for comfort, or too far to care? Finding humor in distant tragedies and close mishaps. Psychol Sci 23(10):1215–1223. doi:10.1177/0956797612443831
13. Berk LS, Tan SA, Berk D (2008) Cortisol and Catecholamine stress hormone decrease is associated with the behavior of perceptual anticipation of mirthful laughter. FASEB J 22(946):11
14. Kounios J, Beeman M (2009) The Aha! Moment the cognitive neuroscience of insight. Curr Dir Psychol Sci 18(4):210–216. doi:10.1111/j.1467-8721.2009.01638.x
15. Bergson H (2003) Laughter: an essay on the meaning of the comic. [Internet] Project Gutenberg [cited 2015/2/7]. <http://www.gutenberg.org/files/4352/4352-h/4352-h.htm>
16. Cramer H, Mentis H, Fernaeus Y (2010) Serious work on playful experiences: a preliminary set of challenges. ‘Fun, seriously?’ In: Proceedings of the workshop at CSCW
17. Shneiderman B (2004) Designing for fun: how can we design user interfaces to be more fun? Interactions 11(5):48–50
18. Venkatesh V (2000) Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. Inf Syst Res 11:342–365. doi:10.1287/isre.11.4.342.11872
19. Yi MY, Hwang Y (2003) Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. Int J Hum Comput Stud 59(4):431–449. doi:10.1016/S1071-5819(03)00114-9
20. Hatfield E, Cacioppo JL, Rapson RL (1993) Emotional contagion. Curr Dir Psychol Sci 2:96–99
21. Vilaythong A, Arnau R, Rosen D, Mascaró N (2006) Humor and hope: can humor increase hope? Humor 16(1):79–89. doi:10.1515/humr.2003.006
22. Hampes WP (2010) The relation between humor styles and empathy. Eur J Psychol 6(3):34–45
23. wit (2010) [Internet]. Random House Kernerman Webster’s College Dictionary, Random House, Inc. [cited November 12 2014]. <http://www.thefreedictionary.com/wit>
24. Boyd B (2004) Laughter and literature: a play theory of humor. Philos Lit 28(1):1–22
25. Heerink M, Krse BJA, Wielinga BJ, Evers V (2008) Enjoyment, intention to use and actual use of a conversational robot by elderly people. In: Proceedings of the ACM/IEEE international conference on human–robot interaction (HRI), pp 113–119. doi:10.1145/1349822.1349838
26. Nishio S, Taura K, Sumioka H, Ishiguro H (2013) Teleoperated android robot as emotion regulation media. Int J Soc Robot 5(4):563–573
27. Hayashi K, Kanda T, Miyashita T, Ishiguro H, Hagita N (2008) Robot manzai: robot conversation as a passive-social medium. Int J HR 5(1):67–86
28. Kahn PH, Ruckert JH, Kanda T, Ishiguro H, Gary HE, Shen S (2014) No joking aside using humor to establish sociality. In: Proceedings of the ACM/IEEE international conference on human–robot interaction (HRI). doi:10.1145/2559636.2559813
29. Lindblom J, Ziemke T (2006) The social body in motion: cognitive development in infants and androids. Connect Sci 18(4):333–346. doi:10.1080/09540090600868888
30. Csikszentmihalyi M, Rathunde K (1993) The measurement of flow in everyday life: towards a theory of emergent motivation. In: Jacobs JE (ed) Developmental perspectives on motivation. Nebraska symposium on motivation. University of Nebraska Press, Lincoln
31. Saerbeck M, Bartneck C (2010) Attribution of affect to robot motion. In: Proceedings of the 5th ACM/IEEE international conference on human–robot interaction (HRI), pp. 53–60
32. Heider F, Simmel M (1944) An experimental study of apparent behavior. Am J Psychol 57(2):243. doi:10.2307/1416950
33. Atkinson R, Hinton D (1992) Funny Business, Episode 1—aired November 22 1992 in UK, Tiger Television Productions, 5:25-6:25, retrieved 2016-10-29 from <https://www.youtube.com/watch?v=E9fsn6lQBV4&list=PL0C4B2B6CC3BCDCCC>
34. Cooney MD, Zanlungo F, Nishio S, Ishiguro H (2012) Designing a flying humanoid robot (FHR): effects of flight on interactive communication. In: Proceedings of the 21st IEEE international symposium on robot and human interactive communication (RO-MAN), pp 364–371. doi:10.1109/ROMAN.2012.6343780
35. Cooney MD, Karlsson SM (2015) Impressions of size-changing in a companion robot. In: Proceedings of the 2nd international conference on physiological computing systems (PhyCS)
36. Goetz J, Kiesler S, Powers A (2003) Matching robot appearance and behavior to tasks to improve human–robot cooperation. In: Proceedings of the workshop on robot and human interactive communication (RO-MAN)
37. Becker-Asano C, Ishiguro H (2009) Laughter in social robotics no laughing matter. In: Proceedings of 8th international workshop on social intelligence design
38. Dragan AD, Holladay R, Srinivasa SS (2014) An analysis of deceptive robot motion. In: Proceedings of robotics: science and systems (RSS)
39. Litoiu A, Ullman D, Kim J, Scassellati B (2015) Evidence that robots trigger a cheating detector in humans. In: Proceedings of the ACM/IEEE international conference on human–robot interaction (HRI), pp 165–172

40. Ng WS, Sharlin E (2011) Collocated interaction with flying robots. In: Proceedings of the IEEE international symposium on robot and human interactive communication (RO-MAN), pp 143–149
41. Cooney MD, Kanda T, Alissandrakis A, Ishiguro H (2013) Designing enjoyable motion-based play interactions with a small humanoid robot. *Int J Soc Robot* 6(2):173–193. doi:[10.1007/s12369-013-0212-0](https://doi.org/10.1007/s12369-013-0212-0)
42. Kanda T, Sato R, Saiwaki N, Ishiguro H (2007) A two-month field trial in an elementary school for long-term human–robot interaction. *IEEE Trans Robot* 23(5):962–971
43. Satake S, Kanda T, Glas DF, Imai M, Ishiguro H, Hagita N (2009) How to approach humans? Strategies for social robots to initiate interaction. In: Proceedings of the ACM/IEEE international conference on human–robot interaction (HRI), pp 109–116
44. Lee MK, Kiesler S, Forlizzi J, Srinivasa S, Rybski P (2010) Gracefully mitigating breakdowns in robotic services. In: Proceedings of the ACM/IEEE international conference on human–robot interaction (HRI), pp 203–210
45. Hobbes T (1840) Human nature. In: Molesworth W (ed) The english works of Thomas Hobbes of Malmesbury, vol IV. Bohn, London
46. Conrad P (1997) It's boring: notes on the meanings of boredom in everyday life. *Qual Sociol* 20(4):465–475
47. Lfdahl A (2005) Preschool teachers' conceptions of children's chaotic play. In: McMahon FF, Lytle DE, Sutton-Smith B (eds) Play: an interdisciplinary synthesis. Play and culture studies, vol 6. University Press of America, Lanham, pp 195–204
48. Zillmann D (2007) Excitation-transfer theory. In: Encyclopedia of children, adolescents, and the media, pp 312–313. doi:[10.4135/9781412952606.n157](https://doi.org/10.4135/9781412952606.n157)
49. Aristotle (1932) Poetics, 1449b in Aristotle. Aristotle in 23 Volumes, vol 23, translated by W.H. Fyfe. Cambridge, MA, Harvard University Press; London, William Heinemann Ltd. [Internet] 1932. [cited 2015/7/16]. <http://www.perseus.tufts.edu/hopper/text?doc=urn:cts:greekLit:tlg0086.tlg034.perseus-eng1:1449b>
50. Burghardt GM (2005) The genesis of animal play: testing the limits. The MIT Press, London
51. Enquist M, Arak A (1994) Symmetry, beauty, and evolution. *Nature* 372(6502):169–172. doi:[10.1038/372169a0](https://doi.org/10.1038/372169a0)
52. Vollmer AL, Muehlig M, Steil JJ, Pitsch K, Fritsch J, Rohlfing KJ et al (2014) Robots show us how to teach them: feedback from robots shapes tutoring behavior during action learning. *PLoS ONE* 9(3):e91349. doi:[10.1371/journal.pone.0091349](https://doi.org/10.1371/journal.pone.0091349)
53. Holladay RM, Dragan AD, Srinivasa SS (2014) Legible robot pointing. In: Proceedings of the international symposium on human and robot communication (Ro-Man). doi:[10.1109/ROMAN.2014.6926256](https://doi.org/10.1109/ROMAN.2014.6926256)
54. Cooney M, Nishio S, Ishiguro H (2014) Affectionate interaction with a small humanoid robot capable of recognizing social touch behavior. *ACM Tran Interact Intell Syst (TiiS)* 4(4):1–32. doi:[10.1145/2685395](https://doi.org/10.1145/2685395)
55. Zucker M, Ratliff N, Dragan AD, Pivtoraiko M, Klingensmith M, Dellin CM et al (2013) CHOMP: covariant hamiltonian optimization for motion planning. *Int J Robot Res* 32(9–10):1164–1193. doi:[10.1177/0278364913488805](https://doi.org/10.1177/0278364913488805)
56. van Someren MW, Barnard YF, Sandberg JAC (1994) The think aloud method: a practical guide to modelling cognitive processes. Academic Press, San Diego
57. Papastefanou G (2013) Reliability and validity of RTR measurement device. *GESIS—Leibniz-Institut fr Sozialwissenschaften*, pp 1–26
58. Drury JL, Keyes B, Yanco HA (2004) LASSOing HRI: analyzing situation awareness in map-centric and video-centric interfaces. In: Proceedings of the ACM/IEEE international conference on human–robot interaction (HRI), pp 279–286
59. Kidd CD, Breazeal C (2004) Effect of a robot on user perceptions. In: Proceedings of 2004 IEEE/IRSI international conference on intelligent robots and systems (IROS), pp 3559–3564
60. Tanaka F, Cicourel A, Movellan JR (2007) Socialization between toddlers and robots at an early childhood education center. In: Proceedings of the National Academy of Sciences of the USA (PNAS), pp 17954–17958. doi:[10.1073/pnas.0707769104](https://doi.org/10.1073/pnas.0707769104)
61. Liu C, Ishi CT, Ishiguro H, Hagita N (2013) Generation of nodding, head tilting and gazing for human–robot speech interaction. *Int J HR* 10(1):1350009-1-19
62. Woods SN, Walters ML, Koay KL, Dautenhahn K (2006) Methodological issues in HRI: a comparison of live and video-based methods in robot to human approach direction trials. In: The 15th IEEE international symposium on robot and human interactive communication (RO-MAN06), pp 51–58
63. Ambady N, Rosenthal R (1992) Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis. *Psychol Bull* 111:256–274
64. Johansson G (1973) Visual perception of biological motion and a model for its analysis. *Percept Psychophys* 14(2):201–211. doi:[10.3758/BF03212378](https://doi.org/10.3758/BF03212378)
65. Veto P, Thill S, Hemen P (2013) Incidental and non-incidental processing of biological motion: orientation, attention and life detection. In: Cooperative minds: social interaction and group dynamics: Proceedings of the 35th annual meeting of the cognitive science society, pp 1528–1533
66. Clark B (2015) How to use the 'Rule of Three' to create engaging content. Copyblogger. Accessed 2016/10/26 at <http://www.copyblogger.com/rule-of-three/>
67. Vargha K, Litovkina AT (2013) Punning in Hungarian anti-proverbs. *Eur J Humour Res* 1(3):15–25
68. Scherer KR (2005) What are emotions? And how can they be measured? *Social Sci Inf* 44(4):695–729. doi:[10.1177/0539018405058216](https://doi.org/10.1177/0539018405058216)
69. Lee MK, Forlizzi J, Kiesler S, Rybski P, Antanitis J, Savetsila S (2012) Personalization in HRI: a longitudinal field experiment. In: Proceedings of the ACM/IEEE international conference on human–robot interaction (HRI), pp 1–8
70. Gockley R, Bruce A, Forlizzi J, Michalowski M, Mundell A, Rosenthal S et al (2005) Designing robots for long-term social interaction. In: Proceedings of the IEEE/IRSI international conference on intelligent robots and systems (IROS), pp 1338–1343
71. Lowe R, Ziemke T (2013) The role of reinforcement in affective computation. In: Proceedings of the conference on computational intelligence for creativity and affective computing (CICAC), pp 17–24. doi:[10.1109/CICAC.2013.6595216](https://doi.org/10.1109/CICAC.2013.6595216)

Martin Cooney received the M.Eng. Degree and Ph.D. degree in systems innovation, engineering science, from Osaka University, Osaka, Japan, in 2011 and 2014. From 2008 to 2013, he was also an Intern Researcher at Advanced Telecommunications Research Institute International (ATR), Kyoto. Since 2014, he has been a researcher at the Intelligent Systems Lab at Halmstad University, Sweden. His research interests include intelligent robotics and human–robot interaction.

Anita Sant'Anna received a B.Eng. degree in Electrical Engineering at Universidade Federal de Santa Catarina in 2007 and a Ph.D. degree in Signals and Systems at Halmstad University, Sweden, in 2012. Since 2013, she has been working as an assistant professor at Halmstad University. Her research interests include signal processing, data analysis, mHealth, and human movement analysis.