Sketching CuddleBits: Coupled Prototyping of Body and Behaviour for an Affective Robot Pet

Paul Bucci[1], Xi Laura Cang[1], Anasazi Valair[1], David Marino[1], Lucia Tseng[1], Merel Jung[2], Jussi Rantala[3], Oliver S. Schneider[1], Karon E. MacLean[1]

[1]University of British Columbia
Vancouver, Canada
{pbucci, maclean}@cs.ubc.ca

[2]University of Twente
Enschede, Netherlands
m.m.jung@utwente.nl

[3]University of Tampere
Tampere, Finland
jussi.e.rantala@sis.uta.fi

ABSTRACT
Social robots that physically display emotion invite natural communication with their human interlocutors, enabling applications like robot-assisted therapy where a complex robot's breathing influences human emotional and physiological state. Using DIY fabrication and assembly, we explore how simple 1-DOF robots can express affect with economy and user customizability, leveraging open-source designs.

We developed low-cost techniques for coupled iteration of a simple robot's body and behaviour, and evaluated its potential to display emotion. Through two user studies, we (1) validated these CuddleBits' ability to express emotions (N=20); (2) sourced a corpus of 72 robot emotion behaviours from participants (N=10); and (3) analyzed it to link underlying parameters to emotional perception (N=14).

We found that CuddleBits can express arousal (activation), and to a lesser degree valence (pleasantness). We also show how a sketch-refine paradigm combined with DIY fabrication and novel input methods enable parametric design of physical emotion display, and discuss how mastering this parsimonious case can give insight into layering simple behaviours in more complex robots.

ACM Classification Keywords
H.5.2 User Interfaces: Prototyping; H.5.2 User Interfaces: Haptic I/O

Author Keywords
Affective computing; human-robot interaction (HRI); Do-It-Yourself (DIY); haptics; physical prototyping.

INTRODUCTION
Emotionally communicative machines are here. Socially expressive robots have potential in therapy, assistance and entertainment; multi-degree-of-freedom (DOF) robots like Probo, Huggable and the Haptic Creature have already demonstrated that discriminable emotions can be rendered with visible or salient haptic motion [27, 33, 38].

If even 1-DOF robots can display emotional content, we could reduce cost, enable expressive Do-It-Yourself (DIY) robots, and inform more complex robot designs. This seems plausible: humans have a powerful ability to anthropomorphize, easily constructing narratives and ascribing complex emotions to non-human entities [13]—even a 1-DOF rod in motion [12].

In this paper, we evaluate the expressive potential for simple 1-DOF robots, and explore the design and recognition of rendered emotional breathing-like behaviours through 1-DOF zoomorphic social robots collectively called CuddleBits.

Motivating our exploration of breathing-like behaviours is a body of work on affective breathing behaviours for both humans and animal-like robots [3, 30]. Although there are many use cases for a robot capable of physical emotion display, we focus here on an affective robot pet [39]. As motivation, consider some scenarios illustrating how a simple robot that can mirror a user's affective state might be emotionally supportive:

1. **Emotion regulation.** Both the user and the robot are initially *Stressed*; the act of calming the robot also calms the user. Steady physical robot breathing has been shown to calm users [30].
2. **Emotion initiation.** Both the user and the robot are initially *Depressed*. The user is told to get the robot to act excited, and as she succeeds, both become *Excited*.
3. **Emotion enhancement.** Both user and robot are *Relaxed*. As they rest together, they both become *more* relaxed.

For a robot to mirror user emotion, it would be useful to automatically generate emotional behaviours that vary across emotional dimensions of arousal (activation) and valence (pleasantness) [25] (Figure 3). However, it is not obvious how these behaviours should look and feel. Even creating apparently simple breathing-like behaviours is complex: parameters of rate, amplitude and waveform must be chosen in keeping with the robot's body; conveying valence may require time variance.

Objectives, Approach and Contributions
In this work, we undertook to understand this new design space. First, we needed to identify tool and metaphor require-

CHI 2017, May 06 - 11, 2017, Denver, CO, USA

ACM 978-1-4503-4655-9/17/05...$15.00
DOI: http://dx.doi.org/10.1145/3025453.3025774

ments for designing emotional robot behaviours in a 1-DOF robot. For example, can a traditional keyframe editor fully support behaviour design? How can we support improvisation for expressive capacity, behaviours, and form factor?

It was important to determine if behaviours for a 1-DOF robot can be designed and consistently distinguished in the basic emotion dimensions of arousal and valence; and which design and engineering parameters give the highest control over these influential perceptual factors. Finally, we wanted to consider difficult-to-control factors and evaluative economy. How do form factor and other non-behavioural factors influence emotion perception? Does seeing the robot live accord the same emotion ratings as seeing it on video?

We devised a rapid prototyping approach that allowed us to design and quickly iterate on a small family of 1-DOF robot bodies (Figure 1). The robots were designed *with sketching/modification as a primary requirement*, thus allowing for iteration on the order of under a minute to less than a day. Simultaneously, we explored *design tools for robot behaviours*, extending a keyframe-based editor for vibrotactile sensations (Macaron [28]) and a new tool for improvisational haptic sketching with vocal input (Voodle).

We developed, tested and analyzed a set of emotion behaviours. Study participants rated the emotional perception of our behaviour set for agreement with validated emotion words. We analyzed the behaviour set for design parameters and characteristic signal features (mean, median, max, min, variance, total variance, and area under the curve) to determine which aspects of robot motion correlate with arousal and valence. We further examined robot form factor and participant perception of emotion under two sets of two conditions: surface {soft and furry vs. hard and exposed}, and viewing condition {live vs. video} in a mixed experiment design.

We contribute:

DIY designs for a 1-DOF robot with a validated range of emotional expression;

Identification of *relationships* between robot behaviour control parameters, and robot-expressed emotion (as consistently rated by participants).

Demonstration of the capability and potential of a sketch-and-refine design approach whose novelty lies in facilitating joint consideration of form and behaviour for 1-DOF affective robots, and can be extended to more complex displays.

In the following, we survey previous work, motivating our choice of behaviour paradigm and design. We then describe two studies: (1) an exploration of robot form factor and how it may influence emotion perception; and (2) a focused look at the physical or engineering parameters that suggest emotional content. Finally, we discuss our findings on the emotion display capability of minimalistic robots and ground them in implications for future evolution.

RELATED WORK

To enable emotional interaction, machines must be able to both *sense* and saliently *display* emotions [10].

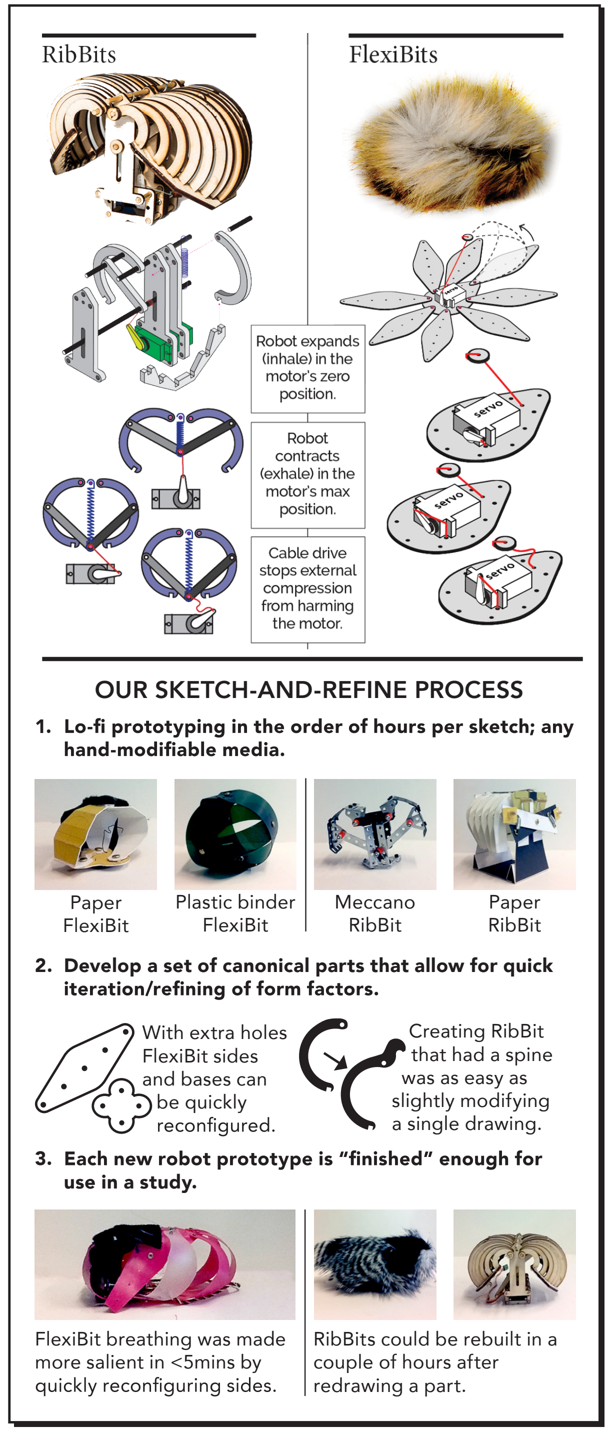

Figure 1. The two CuddleBit families explore very different form factors using similar actuation principles and requirements, but are designed to be quickly modified.

We root our exploration of affective robot motion in the literature, borrowing insight from companion robots' emotional breathing, and considering available emotion models. We take inspiration from other places where physical and behavioural design is coupled, and end with the custom animation methods and tools that underlie our motion rendering process.

Companion Robots, from Complex to Simple

Companion robots that once existed only in science fiction are quickly becoming part of our present reality. Paro, a cute actuated harp seal with soft fur, has been used as a therapy robot in elder care homes to help manage dementia and encourage socialization [34, 35]; its study provides evidence that even simple robot behaviour can produce therapeutic benefits.

The Haptic Creature is a furry mammal-like robot with a multi-DOF haptic emotion display [37], shown to display a variety of emotions with multi-DOF [38]. The physiological effect of one level of sinusoidal robot breathing was studied in a controlled trial, where users held the robot in "breathing" and "non-breathing" conditions. In the first, they experienced a statistically significant decrease in their own heart- and breathing-rates as well as self-reported relaxation [30].

Here we examine the larger *expressive* space of breathing alone. We a) design and validate two distinct robot forms presenting a wide range of affective expression; b) design and validate more complex and nuanced behaviours, introducing control parameters beyond sinusoidal motion variants; and c) achieve these with just 1-DOF breathing motion.

Rendering Emotion Through Breathing

Breathing is a natural expression of emotion [3, 4, 22] that can enhance the recognition of certain affective states, e.g. as displayed on a 3D avatar of a human upper body (including the face) [8]. It is yet unclear how to operationalize design and/or display parameters to distinguish between emotional states. Boiten et al. identify two relevant categories of respiratory parameters: (1) volume and timing parameters (i.e. respiration rate and respiration volume, frequency and amplitude respectively), (2) measures regarding the morphology of the breathing curve (i.e. changes of the breathing curve over time; irregularity) [4]. Our keyframe editor attempts to capture both sets of parameters.

In previous work on similar 1-DOF robots, people petting the small furry bodies reported that their simple periodic motions seemed analogous to breathing, suggesting aliveness [5, 6]. Their behaviours have been likened to other repetitive biological behaviours such as wing motion or heartbeats.

Here, we attempt to access the emotional content of breathing-like motions in a simple pet-like robot (rather than, e.g., a dynamic image of the human torso): specifically, how and to what degree motion parameters signify emotion state.

We also ask whether physical presence, and/or haptic interaction is essential to their emotional expressiveness. In a study of visual proxies of vibrotactile sensations, subjects rated some haptic sensations similarly regardless of whether they experienced it haptically or as a graphical representation [29]. If true here, it may be possible to design haptic emotive robot behaviours using online tools like Mechanical Turk [29].

Emotion Models

For simplicity, we have built our emotion rendering on a conventional two-dimensional emotion model of valence (pleasantness) and arousal (activation) [25]. This model is often discretized by dividing the 2D plane into a grid [26]; and the Positive and Negative Affect Schedule (PANAS) [36] is a set of words with validated mappings to divisions of the resulting affect grid [9]. We use a subset of PANAS words to refer to the grid's extreme corners: *Excited* (positive valence, high arousal), *Relaxed* (positive valence, low arousal), *Depressed* (negative valence, low arousal), and *Stressed* (negative valence, high arousal), as well as others (a total of 20 words) that fall into each quadrant (see Table 1 for the word set used in each study). Participants rated behaviours by indicating how closely a particular behaviour matched a word from the PANAS. Model choice is motivated by dimensionality, with future behaviour generation on continuums of valence and arousal in mind.

Coupled Physical and Behaviour Design

We observe that emotional behaviours are intrinsically bound to the bodies on which they are expressed, and therefore reason that affective robot design will benefit strongly from the flexible, rapid-turnaround physical prototyping methods being popularized by DIY and maker culture.

Our sketch-and-refine process draws on low-fidelity haptic sketching, traditional sculpture and paper craft, rapid prototyping, and animation methods. Our innovation is a process that integrates existing sketching methods by preparing them in forms amenable to quick combination and refinement.

Sketching physical designs

As a design philosophy, *sketching* means quickly generating many versions of an imagined final product. For some engineering problems, detailed specifications can be generated without this step. In contrast, establishing requirements through fast iteration (design-by-prototyping) [16, 18] is typical for those who aim to generate emotional reactions — e.g., industrial designers, set designers and sculptors, who often develop a larger design by making and getting reactions to multiple simultaneous *maquettes* (small physical models).

Our physical design methods are inspired by Moussette's haptic sketching [19, 20], with complex haptic expressions mocked up with low-cost physical media, and by Hoffman and Ju [15], where robots were designed through drawn sketches, prototyping actuated skeletons and 3D modeling.

We take "robot sketching" further. Concerned with haptic as well as visual emotiveness, we prioritize rapid access to user responses to tangible, physical media. Therefore, skipping 3D modelling, we directly implement our robots in low-cost sketch media (wood, plastic) so they are both easy to iterate on and immediately study-ready. We thus design both *for* and *through* sketching: by making on-the-fly modification a primary design requirement, haptic (touchable) robot prototypes can be iterated on within hours or even minutes (Figure 1).

In keeping with Maker ethos, we also target open hardware design, inspired by projects like WoodenHaptics [11].

Animation and tool design

Affective robot control differs from typical robot motion planning, with a goal of communicative rather than functional movement. Affective robot behaviour design consequently draws heavily on puppeting (3D) and animation [24]. Our minimal sketch-and-refine approach stands in contrast to those for higher-DOF affective robots such as Probo [27], where focus on facial expression for emotional display necessitates extensive 3D simulation software to coordinate actuation.

Animators tout the benefits of sketching to develop believable motion parameters where a sketch's subordinated detail presents opportunities to zoom in, problem-solve, then back out — a methodology not lost on haptic designers [21]. Recently developed haptic design tools include Macaron [28], which borrows the animation tenet of keyframes, directly manipulating the vibrotactile sensation to match notable events (or frames) in an analogous animation.

DESIGN METHODS

To validate a behaviour's emotion content, we developed 1-DOF social robots by rapidly iterating form in concert with identification of parameters controlling expressive movement.

Two major form factors emerged (collectively called the 'CuddleBits'): a rigid wood-based form resembling a rib cage ('RibBit') and a flexible, plastic-based form resembling a ball of fur ('FlexiBit').

We iterated on mechanical form until satisfied with the prototypes' tactility and expressive possibilities of movement. We generated dozens of throw-away prototypes from paper, Meccano, plastic and wood (see Figure 1 for an illustration of the process, and select prototypes). After initial lo-fi sketches, a key requirement we set for medium-fi prototypes was to design *for sketching*, i.e. design our robots to be easily modifiable using simple and repeatable parts. By applying a deconstructive, sketch-based process to a hard, subjective design problem, we were able to quickly react and adapt to both control and expressive requirements on the fly.

Considerations

Robust compliance and handling affordances: Our prototypes had to withstand the pressure of a human hand, but we found that the structure's basic pliability and apparent fragility also directed its handler's approach: pliability afforded rough play (squishing, hitting, throwing) while rigidity incited gentleness (holding, cupping, stroking).

Touch: Prototypes needed to convey some kind of biological behaviour; thus, the materials needed to afford playfulness and liveliness.

Believability: We drew on caricature and internal consistency for believability; we took cues from the natural world but situated the CuddleBits in their own genre rather than mimicking a real animal. RibBit's inner mechanics are evident, allowing natural material affordances but cueing user expectations; FlexiBit's soft fur elicited stroking and its limbless form suggested no locomotor ability.

Backend: The Bits are powered and controlled by an Arduino Uno, from a NodeJS server (`nodejs.org/en`) using the Johnny-Five JS robotics framework (`johnny-five.io`). Javascript construction facilitates connection to front-end applications and widely-available web frameworks. Using a single language for front- and back-end facilitates seamless development; using web technologies allows for transparency in widget design (i.e. using the browser 'inspect' tool makes for faster debugging than with Java UI packages). As such, rapidly iterating on control mechanisms was also fast.

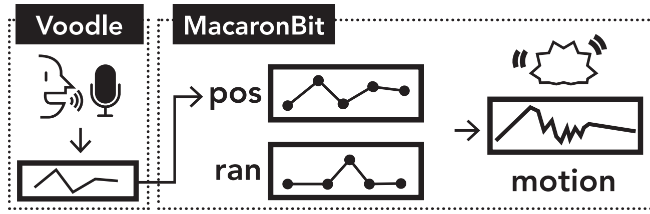

Figure 2. For Study 2a, users had two design tools to help create behaviours: Voodle and MacaronBit. With Voodle, users could optionally sketch a robot motion using their voice. Their vocal input was imported into MacaronBit as raw position keyframes (shown as 'pos' above). Users could modify the waveform by manipulating other parameters, specifically randomness (shown as 'ran'), max and min position. MacaronBit includes standard keyframe editing functions.

Extendable and Open design: The CuddleBits were designed to be easily extended with little effort or expertise. Full CuddleBit source documents are shared: pattern files, code and a manual for assembly and extension[1].

CuddleBit Visual Design Patterns

Modifiable design patterns make varying robot shape quick, from under 2 hours up to 2 days. For each family, we produced many models (Figure 1) and chose the most visually and haptically salient to evaluate. Each pattern includes both the atomic units of construction and the narrative sense conveyed by the aesthetic and material presentation, detailed below.

The FlexiBit: Like a simple sewn sphere, Flexi's ribs are plastic slices fixed to a base like petals of a flower. These are generated by adjusting a stencil, printing and cutting the pattern from plastic sheets with a knife, and joining them with machine screws. Slices are scalable for smaller or larger Bits; to adjust shape, only some slices and/or the base are varied.

Plastic flexibility, volume, and curvature provide passive compliance and natural feel under a faux fur cover. It is often compared to a Tribble (a fuzzy alien species from the Star Trek universe): the plastic frame evokes a compliant torso.

The RibBit: A wooden ribcage on a stand, its rigid actuation gets compliance from internal springs. In counterpoint to Flexi, mechanics are fully exposed, with no attempt at material realism. It comes alive only with suggestive movement. Each rib is laser cut from an easily modifiable digital pattern, and assembled by wood-gluing parts together, with BBQ skewers as pins and rods. Further versions of the RibBit include fur and ridges that form an inflexible 'spine'; these explorations were prototyped in low cost materials (like paper) first.

RibBit has a naturalistic skeletal aesthetic due to its wooden construction. Fitting comfortably in the contours of a hand, the structural rigidity provides notable haptic feedback even when covered with fur.

[1] `www.cs.ubc.ca/labs/spin/cuddlebits`: pattern files, code and a DIY manual for assembly and extension.

Animation Tools

We used internally developed haptic design tools to quickly sketch and refine behaviours: (1) MacaronBit, a keyframe-based vibrotactile effect editor and (2) Voodle, a robot behaviour sketching tool.

Early tool attempts: While extending Macaron for robot behaviour design seems natural in retrospect, at the time it was not obvious to repurpose a vibrotactile design tool for 1-DOF robot motions. A keyframe editor balances the concerns of precision and refinement with the easy sketching abilities afforded through direct manipulation; Macaron provides focused control over a short window (in the order of seconds).

Our group made many attempts at behaviour design/puppeting tools and techniques before MacaronBit and Voodle, including a free-hand vector drawing tool based on `paperjs.org`, a browser-based timeline editor [1]; and for direct position control, a force-feedback knob [31]. The timeline editor was unintuitive; the drawing tool and force-feedback knob required too much fine control for large movements, and did not provide enough control for small movements (i.e. entire behaviour would have to be re-drawn per sketch; human hands cannot move fast enough to express fine motions such as fluttering). MacaronBit allowed for precise control of both large and fine motions, and Voodle allowed for quick sketching and iteration since it gave immediate continuous feedback.

MacaronBit: Macaron is a open-source web-based keyframe editor for designing vibrotactile sensations, using amplitude and frequency of a vibration [28]. As our first pass at designing a robot behaviour editor, we extended Macaron to robot position control, calling the result MacaronBit (Figure 2).

In developing MacaronBit, we started with a pure sine-wave and adjusted its parameters: frequency, amplitude, bias, and amplitude/frequency variability. Its support of immediate playback, key-framing (parameter interpolation between key points), waveform generation, and click-and-drag editing sped up iteration. For participant-designed behaviours, we switched to direct position control, where design parameters were position, randomness, and max and min position (Figure 2).

***Voodle** – vocal doodling for sketching behaviours:* Although traditional keyframe behaviour design has strong roots in animation, it is somewhat unintuitive for non-experts. To better afford improvisation and sketching, intrinsic to the process of design, we attempted a number of sketching tools ranging from direct puppeting with a haptic knob, to a vector-based drawing tool, to a vocal sketching tool. The last, Voodle, translates sound into robot motion by mapping vocal amplitude and pitch in a linear combination to servo position, and proved to be the most intuitive. For example, if you were to laugh, the robot would convulse in time to your laughter. By adjusting a set of parameters to tune the voice-to-motion mapping, one could capture different vocal features to be expressed on the robot, i.e. shifting the 'smoothness' of the motion, changing the influence of pitch or amplitude in the output motion.

STUDY METHODS

We ran two distinct studies: Study 1 is an initial look into robot expressive capacity and is further expanded on in Study 2. For both, we generated a set of behaviours to determine emotional recognizability via participant ratings. In Study 1, behaviours were designed by researchers, rated by participants (N=20); in Study 2a, behaviours were designed by dyads consisting of a naïve participant (N=10) and an expert animator. Participants in Study 2b (N=14) rated behaviours designed in Study 2a under two viewing conditions: behaviours viewed live and via video. In all cases, we avoided referring to behaviours as 'breathing', allowing participants to interpret for themselves.

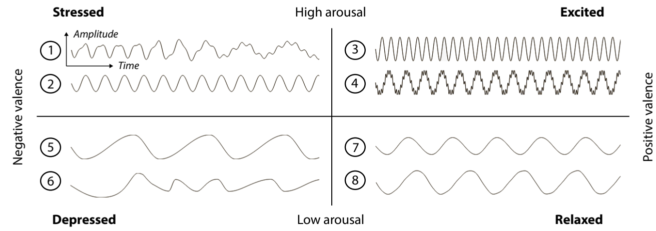

Figure 3. Waveforms of Study 1 behaviours as designed by researchers. Each quadrant is represented by a PANAS affect word corresponding to the extremes along (valence, arousal) axes, i.e. *Excited* is high-arousal, positive-valence.

Study 1 - Robot Form Factor and Behaviour Display

We evaluated the emotional expression capabilities of our two CuddleBit forms (FlexiBit and RibBit) on eight behaviours representing four emotional states. Specifically, we asked:

RQ1: Can 1-DOF robot movements be perceived as communicating different valence and arousal states?
Hypothesis: Different levels of arousal will be interpreted more accurately than different levels of valence.

RQ2: How is interpretation of emotional content influenced by robot materiality, e.g., a soft furry texture?
Hypothesis: FlexiBit's behaviour will be perceived as conveying more positive valence than RibBit's.

Development of Study 1 behaviours

Team members created and agreed upon two breathing behaviours for each quadrant of the affective grid [26]: *Depressed*, *Excited*, *Relaxed*, or *Stressed*, for a total of 8 behaviours (represented as motion waveforms in Figure 3). Each emotion word typifies the extreme of its emotion quadrant (i.e. *Stressed* is high-arousal, negative-valence).

Study 1 Methods: body and behaviour evaluation

Participants: 20 participants, aged 20–40 with cultural backgrounds from North America, Europe, Southeast Asia, Middle East and Africa, were compensated $5 for 30 minute sessions.

Procedure: Participants were given the task of rating each behaviour on a 5-point semantic differential (−2 Mismatch to +2 Match) for two different robots displaying four emotions: *Depressed*, *Excited*, *Relaxed*, or *Stressed*. For instance, for "FlexiBit feels stressed", a participant would play each behaviour and rate how well it matched the robot portraying stress. During playback and rating, participants kept one hand on the robot, and moused with the other; motion was experienced largely haptically. Noise-cancelling headphones played pink noise to mask mechanical noises; instructions were communicated by microphone.

Ratings for each robot were performed separately. Robot block order was counterbalanced, with an enforced 2m rest. For each block, all four emotions were presented on the same screen so participants could compare globally. Behaviours (15s clips) could be played at will during the block. Order of behaviours and emotion was randomised by participant. To reduce cognitive load, participants saw the same behaviour/emotion order for the second block. In total, each participant performed 64 ratings (8 behaviours × 4 emotions × 2 robots). Afterwards, a semi-structured interview was conducted.

Study 1 Results

We compared ratings of each pair of behaviours designed for the same emotion word with a pairwise Wilcoxon signed-rank tests with Bonferroni correction (Figure 4). Ratings of the two designed behaviours for the same emotion quadrant were not significantly different ($\alpha = .050/8 = .006$; all p's $\geq .059$). Thus, we averaged ratings into four pairs by emotion target (e.g., (1) & (2) in Figure 3).

Effect of emotion quadrant on behaviour ratings (significant). Friedman's test on behaviour ratings showed significant differences between behaviours per emotion for both robots (all p's $< .001$). Post hoc analyses using Wilcoxon signed-rank tests were conducted with a Bonferroni correction ($\alpha = .050/6 = .008$) to further analyse the effect of **emotion condition** on *researcher-designed behaviours*:

– **Stressed**, **Excited**, or **Relaxed**: There were significant differences between high and low arousal behaviours (*Stressed-Depressed*, *Stressed-Relaxed*, *Excited-Depressed* and *Excited-Relaxed*, all p's $\leq .002$); but none between behaviours with the same arousal level but different valence content.

Effect of robot on behaviour ratings (not significant). Wilcoxon signed-rank tests with Bonferroni correction showed no statistically significant differences between ratings of emotions displayed on the two distinct robot forms ($\alpha = .050/16 = .003$; all p's $\geq .026$).

Duration (not significant). A two-way (2 robots × 4 emotions) repeated measures ANOVA showed no significant differences in the time spent on rating behaviours (all p's $\geq .079$), suggesting each emotion rating was undertaken with similar care.

Study 1 Takeaways

Hypothesis 1: Different levels of arousal are easier to interpret than different levels of valence. – *Supported.*

In general, participants were able to perceive differences in behaviours designed to convey high or low arousal. Speed or frequency was most mentioned for arousal variation: low arousal from low frequency and high arousal from high frequency. Participants found interpreting valence more difficult. Thus, behaviours on this 1-DOF display corroborates earlier findings in regards to both dimensions [7, 23, 39].

We posit that the difficulties in determining valence may be due in part to the restrictive range of behaviours. All designs were based on the perception and imagination of three computer science researchers, which may not be broadly generalizable as effective emotional displays.

Improvement: Behaviours may have more range or discernible valence when sourced from a more diverse group of designers. To increase emotional variance in Study 2, *we recruited participants (N=10), the majority of whom were employed in creative roles to create the behaviours with an expert designer*. Participants were encouraged to puppet robot movements, act out desired movements, and interact with the robot until they were satisfied with the emotional displays.

Hypothesis 2: FlexiBit's behaviour will be perceived as conveying more positive valence than RibBit's. – *Not supported.*

In post-study interviews, participants reported the movement expressed by the two robot forms as sensorially but not necessarily emotionally different. FlexiBit felt nicer to touch, but its motion was less precise. RibBit's movements were interpreted as breathing or a heartbeat despite the exposed inner workings emphasizing the 'machine-ness' of the robot.

Unexpectedly, while participants specified preferences for FlexiBit's fur and RibBit's motor precision, pairwise comparisons of the same emotions revealed no significant difference between robots. Movement rather than materiality dominated how participants interpreted emotional expression; although visual access to form was restricted during movement, tactility might have modulated perception of, e.g., life-likeness.

Improvement: Whereas robot form factor had little to no influence on emotion recognition results, it did influence how participants perceived the robot. We selected characteristics to emphasize for a second round of robot prototyping, *producing a new robot for Study 2*. We focused on characteristics that participants referenced as salient or pleasing in interviews, such as fur, texture, and body firmness.

Starting from paper prototypes, we iterated on the RibBit form factor to increase haptic salience and to incorporate positive FlexiBit features. After exploring bumps on the ribs, spine configuration, fur textures, and rib count, we converged on a form that had fewer ribs, dense fur, and a prominent spine. This combined the favourite features of the RibBit (crisp motion and haptic feedback) with the FlexiBit's cuddliness. With rapid prototyping methods, each paper/lo-fi sketch could be explored in less than an hour; full new robot prototypes took about two hours to modify design files, half an hour to laser cut, and about two hours to assemble.

STUDY 2

Study 2 consists of two parts with two separate groups of participants. In Study 2a (N=10), we partnered naïve users with an expert animator (an author) to produce an interesting and diverse set of behaviours, generating a library of emotional robot behaviours labeled by affective quadrant. In Study 2b (N=14), participants rated the behaviours generated in Study 2a in terms of PANAS words.

Study 2a – Generating behaviours with participants

Study 2a Methods: behaviour generation

Participants: We recruited 10 naïve participants (6 were pet owners). We directly recruited participants with artistic and/or

Figure 4. Mean behaviour ratings (+2 for Match; -2 for Not Match) for FlexiBit grouped by the researcher-designed behaviours (horizontal) and the emotion word against which participants rated behaviours (vertical). Researcher-designed behaviours correspond with (1) to (8) in Figure 3. RibBit scores were similar and omitted for space.

Unpleasant Activated	Pleasant Deactivated	Pleasant Activated	Unpleasant Deactivated
stressed† *	relaxed†‡*	excited†‡*	depressed†*
upset‡*	calm‡*	attentive‡*	drowsy‡*
scared‡*	at rest‡*	determined‡*	bored‡*
guilty‡	serene‡	proud‡	dull‡
hostile‡	at ease‡	enthusiastic‡	sluggish‡
nervous‡			droopy‡

Table 1. Affect Grid quadrants of PANAS emotion words. † represents words used in Study 1; ‡represents words used in Study 2a; * represents words used in Study 2b.

performance experience (7/10) as they were likely to be more comfortable with improvisation and design. All were compensated $10 per session.

Procedure: Participants were introduced to Voodle and MacaronBit. After a warm-up exercise, they were given a PANAS word representing each quadrant of the affect grid (see Table 1 for word list) and asked to sketch the emotion behaviour using Voodle. Once they were satisfied with their sketch, it was imported into MacaronBit where they worked with the expert animator to refine their design. This was repeated for all tasks; participants could decline to use either interface if they felt that the behaviour design did not require it.

In total, each participant was given four words (one per affect quadrant). Words were taken randomly from each quadrant, as was order of quadrant presentation. Each session took ∼60 mins, ending with a semi-structured interview.

Study 2a Results

Voodle was used beforehand to sketch behaviours in most cases. When a participant had a clear idea of what the behaviour should look like, both sketching and refining was performed in MacaronBit.

A library of 72 behaviours labelled by emotion word was generated (participants designed multiple behaviours per word); analysis of these behaviours presented in Study 2b results.

Study 2b - Rating Behaviours Live vs Video

Of the 72–item behaviour set generated in Study 2a, Study 2b used a subset of 16: five researchers selected the most representative designs, converging on the top four per quadrant. Under two viewing conditions {live, video}, participants chose three words that best represented the displayed behaviours and rated their confidence in each chosen word, as well as one or more words that least represented the behaviour. Participants rated words ahead of time in terms of arousal and valence. Ratings per participant and per viewing condition were combined into a single (valence, arousal) point (described below). Through this, we explored the following:

RQ3: Is there a difference in viewing conditions?
Hypothesis: Participants will rate behaviours similarly regardless of viewing condition.

RQ4: Are behaviours consistently distinguishable?
Hypothesis: Each behaviour will be distinguishable.

RQ5: Which behaviour design and waveform features correlate with rated dimensions of arousal and valence?
Conjecture: Features that are characteristic of variability will correlate with valence, while features that are characteristic of speed will correlate with arousal.

Study 2b Methods: behaviour evaluation

Participants: We recruited 14 naïve participants (4 male), aged 22–35. 12 participants were fully proficient in English; the remaining 2 had advanced working knowledge. Out of 14 participants, 13 reported having at least some interaction with pets; 6 rarely interacted with robots, and 8 never interacted with robots. All were compensated $15 per session.

Procedure: Participants were seated, introduced to a fur-covered RibBit, and asked to touch the robot to reduce novelty effects. To calibrate emotion words, participants rated the valence and arousal of 12 words on a 9-point scale (Table 1). Participants then viewed the 16 robot behaviours in two counterbalanced viewing condition blocks {live, video}.

In the live condition, participants could physically interact with the CuddleBit while playing each robot behaviour via MacaronBit. Noise-cancelling headphones played pink noise to mask robot noise. In the video condition, participants watched silent videos of the CuddleBit performing the same behaviours (side view, 640x360 px, 30fps). In both conditions, behaviour order was randomized for each participant.

In each viewing condition, participants were asked to choose 3 emotion words that best represented the behaviour from a list of the 12 emotion words they calibrated previously, indicating their confidence level of each word on a 5-point Likert scale. They watched 16 behaviours and answered qualitative follow-up questions. After an optional 5 minute break, this process was repeated, with condition block counterbalanced. Including a semi-structured interview, the session took ∼60 minutes.

Study 2b data preprocessing

Before each session, participants calibrated the emotion words that they would be using by rating each in terms of arousal and valence. Using the calibrated list of emotion words, we constructed vectors of (v = valence, a = arousal) for each word,

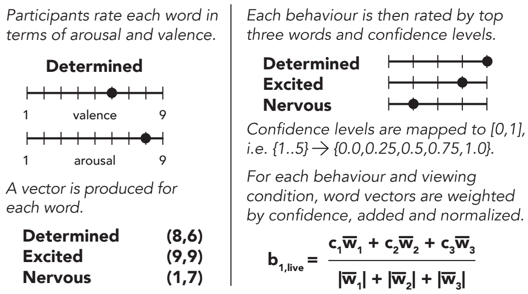

Figure 5. For each behaviour and viewing condition, a single vector was calculated by adding the vectors of the top three words that participants chose, weighted by confidence levels. Word vectors were determined at the beginning of the session, when participants rated each word in terms of arousal and valence.

where $1 < v, a < 9$. For each behaviour and viewing condition, the best three words were weighted by their confidence values, added and normalized. This produced a single vector of (v, a) for each behaviour and viewing condition (Figure 5).

Study 2b Data verification

Before the following analysis, we ran a series of data verifications to ensure consistency in each participant's responses.

Due to the high subjectivity of the kinds of emotions people will associate with different words, the participant-calibrated emotion words were checked for consistency with the expected PANAS quadrants. For all participants, no more than two words disagreed with the PANAS quadrants; as such, we took the participant rated words to be reasonably calibrated.

Similarly, for each behaviour per view condition, the best three rated words were checked both against themselves, and against the selected least representative word(s). Roughly 50 per cent agreed within a reasonable margin of error across either valence or arousal; 30 per cent agreed across both valence and arousal; 20 per cent either did not agree or were inconclusive.

To determine whether our confidence value weighting scheme was valid, we performed both a visual inspection of word distribution and confusion matrices with design labels. With no weighting scheme, data was heavily biased towards (positive valence, high arousal) ratings, which did not agree with our qualitative results or a reasonable reading of our quantitative results. As such, a linear weighting scheme was determined to be the least biased, such that confidence ratings of $\{1,2,3,4,5\}$ were mapped to $\{0.0, 0.25, 0.5, 0.75, 1.0\}$.

Study 2b Analysis and Results

We summarize our findings from Study 2b. All significant results are reported at $p < .05$ level of significance.

RQ3: Is there a difference in viewing conditions?

Hypothesis: Participants will rate behaviours similarly regardless of viewing condition. – *Not supported.*

Behaviour label $\times$ **Viewing condition**: We found a significant effect for viewing condition ($Pillai = 0.563$, $F(2,415) = 6.87$) and behaviour label (0.563, $F(30,832) = 10.86$). We did not find an interaction effect ($p = .33$). Although there is evidence to suggest that participants do rate behaviours differently, since they also rate viewing conditions differently, we should be careful in using video as a proxy for live robot behaviour display.

Behaviour label quadrant $\times$ **Viewing condition**: We found a significant effect for viewing condition ($Pillai = 0.441$, $F(6,880) = 41.43$) and by collecting designs by quadrant (*e.g. Hostile* and *Upset* are both high-arousal, negative-valence emotions), ($Pillai = 0.030$, $F(2,439) = 8.705$), and the interaction effect.

Duration: Through 2-way ANOVAs, we found significance in duration between viewing conditions wherein participants took longer to rate live ($\mu = 72.49s$, $\sigma = 40.69s$) than via video ($\mu = 64.13s$, $\sigma = 29.28s$) per behaviour, corroborated in that live behaviours ($\mu = 2.36$, $\sigma = 1.51$) were played more times than the corresponding video ($\mu = 1.96$, $\sigma = 1.18$). The more time spent on live behaviours could be due to more information conveyed or more interest as participants interpret the motion and/or haptic expression.

RQ4: Are behaviours consistently distinguishable?

Hypothesis: Each behaviour will be consistently distinguishable. – *Partially supported.*

Behaviour label $\times$ **Participant**: As behaviours ($Pillai = 0.917$, $F(30,448) = 12.649$), participants ratings ($Pillai = 0.671$, $F(26,448) = 8.705$), and the interaction are all significant, we determine that the behaviours are distinguishable by participant.

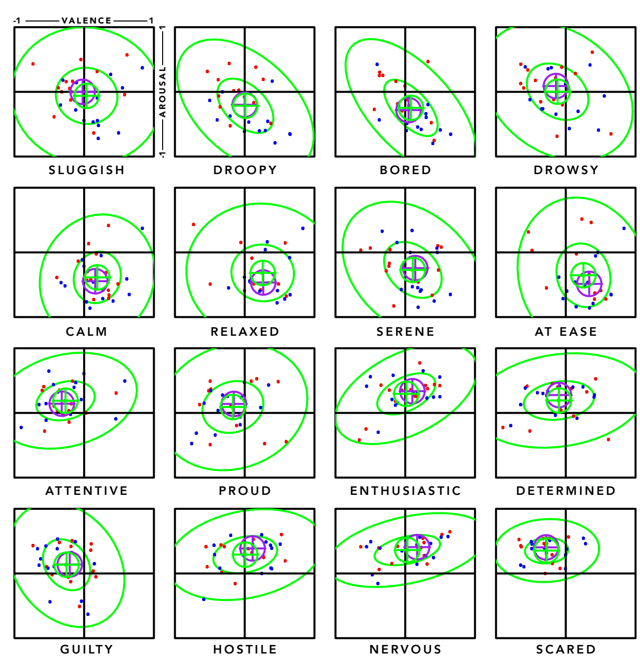

Figure 6. Each plot shows a single behaviour's arousal (-1,1) and valence (-1,1) ratings. Live viewing condition is in red, video in blue. Green ellipses show confidence intervals at 5% and 97.5%. Green cross is mean, purple cross is median. Each plot corresponds to a single PANAS word, each row corresponds to an affect grid quadrant. Rows order from the top: *Depressed*, *Relaxed*, *Excited*, *Stressed*.

Through Figure 6, we examine rating consistency by behaviour and quadrant. Negative-valence, low-arousal (*Depressed*) behaviours have the largest dispersion in rating for both dimensions, suggesting that they are the most difficult for participants to classify. Low-arousal, positive-valence (*Relaxed*) behaviours are more consistently concentrated towards the relaxed quadrants.

Both high-arousal, negative-valence (*Stressed*) and high-arousal, positive-valence (*Excited*) behaviours are concentrated in the high-arousal half, yet highly dispersed across valence, suggesting valence is difficult to determine for certain high-arousal behaviours.

Overall, behaviours designed for a representative quadrant may not necessarily be interpreted as such. *Determined*, for example, was interpreted as negative-valence with high-arousal, a contrast to the intended positive-valence high-arousal. Finally, live behaviours (red in Figure 6) are more dispersed than video behaviours (blue). This illustrates a higher variation in how participants rated live than video behaviours.

RQ5: Which behaviour design and waveform features correlate with rated dimensions of arousal and valence?

Conjecture: Features that are characteristic of variability will correlate with valence, features that are characteristic of speed will correlate with arousal. – *Partially supported*.

Analysis using machine learning techniques was performed as a preliminary step to understand which features might be most relevant. Using the full set of designed behaviours from Study 2a and their associated design labels, we trained a Random Forest classifier on statistical features calculated from design and output waveform attributes. Since each behaviour was output as a waveform, we could decompose the waveform using MacaronBit design parameters, and describe them using keyframe count and standard statistical features (min, max, mean, median, variance, total variance, area under the curve) on keyframe values. The same statistical features were calculated for the output waveform.

Each behaviour label was mapped to the original PANAS quadrant (called here *design quadrant*). When running 20-fold cross-validation classifying on design quadrant, the Random Forest classifier achieved between 66% and 72%, full feature subset for the former, and an optimal subset for the latter (chance=25%). Top performing features were **position**: keyframe count, range, total variance; **random**: max, min.

Note that the selected features are related to waveform complexity. If the *random* parameter was set high, then the waveform would have a high amount of variation. Similarly, if there were a high number of *position* keyframes, the waveform would have a lot of variation.

Feature Selection

A correlation matrix was constructed between arousal and valence for 16 participant-rated behaviours (per viewing condition), and for the 72 participant-and-researcher generated unrated behaviours.

As seen in Figure 7, arousal has stronger correlation within the feature vector than valence. Features with strong positive correlation to arousal are those that also correspond with the widest, fastest, and most erratic motions, such as position keyframe count, position range, and random maximum.

Waveform						Random						Position						
live_a	live_v	video_a	video_v	unrated_a	unrated_v	live_a	live_v	video_a	video_v	unrated_a	unrated_v	live_a	live_v	video_a	video_v	unrated_a	unrated_v	
n/a	n/a	n/a	n/a	n/a	n/a	.02	-.01	-.05	.02	-.03	-.19	.39	-.02	.49	-.13	.39	-.15	numkf
.03	-.06	.09	-.16	-.12	.08	.40	-.10	.46	-.07	.35	-.05	.01	-.16	.02	-.10	-.10	.01	med
-.03	.07	-.14	.11	-.12	-.21	.43	-.11	.45	-.09	.36	-.17	.13	-.10	.16	-.08	.06	-.25	max
-.20	.16	-.15	.17	-.10	.17	.28	-.09	.32	-.06	.28	-.18	-.34	.04	-.35	.09	-.24	.18	min
.26	-.13	.26	-.18	-.01	-.22	.22	-.02	.21	-.02	.23	-.10	-.01	.03	-.04	.12	.05	-.08	var
-.11	.00	-.16	-.04	.18	-.16	.13	-.08	.05	-.04	.20	-.13	.37	.00	.44	-.05	.49	-.08	totvar
.15	-.10	.07	-.10	.06	-.24	.25	-.04	.24	-.06	.30	-.12	.34	-.07	.35	-.11	.24	-.31	range
-.05	.00	-.02	-.07	-.15	.07	.40	-.11	.43	-.08	.36	-.12	-.22	-.08	-.25	.04	-.19	.10	mean
-.05	.00	-.02	-.07	-.15	.07	.29	-.06	.23	-.04	.23	-.04	.39	-.04	.49	-.15	.36	-.15	auc

(Colour scale: correlation, +0.5 to -0.5)

Figure 7. Correlation results from behaviours that were designed for an emotion label but unrated by participants (marked *unrated* above) were calculated on all 72 designs from Study 2a; correlation results from participant-ratings were calculated on the 16 behaviours from Study 2b (marked by viewing condition). A strong positive correlation is shown between the position total variance for all arousal columns ($unrated_a$, $video_a$, $live_a$) – the higher the total variance, the higher the arousal.

Valence has much weaker correlation overall, and particularly low absolute correlation values in the participant-rated analysis. However, within the unrated behaviours, the top correlated features are also indicators of waveform complexity and are negatively correlated with valence, i.e. the more complex a behaviour is, the less it is deemed to be a pleasant behaviour.

Participant Experience

Interviews with participants were audio-recorded, transcribed, and coded for themes and keywords by a single researcher using an affinity diagram and constant comparison. Open-ended written responses from participants for both live and video viewing conditions were analyzed with the same techniques.

In contrast to video, participants emphasized the importance of haptic feedback (7/14), the ability to view the robot from multiple angles (4/14), the increased engagement and accessibility that resulted from the live interaction (5/14); 9/14 touched the robot while playing behaviours. Of these, three reported wanting to touch the robot when they were having difficulty interpreting the behaviour.

> "*Feeling the movements rather than just watching them helped me get a better sense of an interpretation of what the emotion was.*" –P05 (corroborated by P09, P11)

Participants reported that some of the given emotion words were ambiguous (particularly *Depressed*, *Attentive*, *Excited*, and *Stressed*), due to a lack of context and visual cues.

> "*There were some emotions where it was pretty ambiguous. [There are] different connotations depending on exactly what the context is.*" –P03 (P02, P10)

Several participants (4/14) interpreted a combination of emotions while observing the robot behaviours, with emotions happening either simultaneously or sequentially.

> "*I think [the emotions] are happening in order. So sometimes [it's] excited, and then after the stimulus is gone, it becomes bored again.*" –P06 (P04, P07, P09)

Participants' description of their process for labelling robot behaviours included relying on their experience with animal behaviours (4/14), their experience with human emotions (2/14), and interpretation of the "heartbeat" or "breathing" of the robot's movement (7/14). 5/14 participants mentioned that it was difficult to interpret emotions from the robot behaviours.

DISCUSSION

Distinguishing Arousal and Valence

Results show arousal levels are easy to distinguish, but valence is more difficult. This is consistent with many human emotion classification studies, such as with facial expressions: it is difficult to distinguish emotions across valence, even by humans of other humans.

As an example, an enthusiastic soccer player yells after a goal: without knowing which side the soccer player is on, it is difficult to visually distinguish between a yell of anguish or victory [2]. Similarly, for low arousal states, it might be difficult to tell the difference between someone who is relaxed or depressed.

Observers may always need to rely on context for valence, either through extended interaction, or through external environmental or situational cues. Arousal behaviour design might be more important to 'get right' than valence behaviour design, especially if interaction and contextual cues are stronger than any inherent behaviour features. Consider watching a scary movie with a robot; the robot reacts with a high-arousal behaviour. If you assume it is a sympathetic creature, you will likely believe it is reacting with you to the movie. There is also evidence to suggest that valence and arousal are not truly orthogonal, and may co-vary [17].

Even in this restricted context there are clear indicators of validated emotional behaviours. For instance, labels *Guilty* and *Scared* are rated as high arousal-negative valence whereas *Relaxed* and *At Ease* are decidedly low arousal-positive valence, as determined by tight agreement along the axes in Figure 6.

Design parameter features that produce a behaviour with a wider range of motion or more complex motion (i.e. *number of keyframes* for position, *maximum* randomness) correlate with higher arousal and lower valence (Figure 7). Controlling these parameters can create convincing behaviours across emotion dimensions: increased randomness generates negative-valence and high-arousal behaviours as will a larger range of motion.

Form Factor, Viewing Condition and Context

Results from Study 1 suggest that the same behaviours are interpreted similarly on RibBit or FlexiBit. Significant differences between viewing conditions in Study 2b suggest that physical presence does impact a user's emotional understanding of the robot. Video behaviours were rated more consistently than live behaviours which could be due to the robot having a more limited expressive range on video than live; participants reported that the haptic sensations allowed for more nuanced interpretations of displayed emotion. We conjecture that a live robot can express *more*, but is also more open to situational influence, since a live environment is much more sensorially rich than the same environment depicted in video. Difference in live vs. video viewing conditions is reported elsewhere, *e.g.* infants have been shown to pay more attention to live agents than agents on video [14, 32].

Through heuristic evaluations, we found that the narrative considerations of each robot changed with design variations within a single form factor. For example, we were much more inclined to baby a small FlexiBit over the larger one. Yet, despite the contrast in the character of each robot, participants rated behaviours similarly across form factors.

Although there are more behaviour metaphors to explore even in 1-DOF, we acknowledge that more degrees of freedom may improve valence recognition through actuation of body parts for clear emotional signification. For example, if the robot had a head, raising the head might correspond with higher valence; if the robot had a tail, wagging the tail might do the same.

CONCLUSION AND FUTURE WORK

We evaluated the expressive potential for simple 1-DOF robots by exploring the design and recognition of displayed emotional breathing-like behaviours on a 1-DOF zoomorphic social robot called the CuddleBit.

Through a sketch-and-refine process, we developed a series of robot prototypes and animation techniques, a library of affective behaviours, and validated a subset of those behaviours with naïve participants. We improved on one robot design (the RibBit) in a second iteration, incorporating participant feedback to emphasize haptic and emotional salience. We analyzed a set of characteristic features of design parameters, and determined which features correlated with arousal and valence, targeting the generation of affective behaviours.

Our results show promise for further design parameter exploration in producing breathing curve irregularity, supported also by the literature [4]. Future work will examine longer interactions that dynamically vary behaviour parameters, rather than simply displaying pre-recorded motion. To auto-generate *Distress*, the robot could briefly increase waveform randomness and maximum position; then in response to a soft touch, reductions in these parameters may resolve as *Relaxation*.

Future work will address 1-DOF robots with different expressive actuations (e.g. spine or head movement) as well as integrate multiple 1-DOF robots into a multi-DOF design. By fully exploring each DOF in isolation through the same coupled body and behaviour design process, the expressive capacity of each addition can be controlled and iterated on. Layering single DOFs to produce multi-DOF behaviours can filter the design space and reduce late-stage failures. Using the sketch-and-refine process presented here along with behaviour testing techniques, designers should be able to direct explorations based on concrete validations, with a fully-functional, studied prototype produced at each stage.

ACKNOWLEDGMENTS

This publication was supported by NSERC, the Dutch national program COMMIT and the Academy of Finland (Decision #260026). Thanks to Jon Nakane at The UBC EngPhys Project Lab for help; Derek Tan and Lesha Koop at the Beaty Biodiversity Museum for inspiration.

REFERENCES

1. Jeff Allen, Laura Cang, Michael Phan-Ba, Andrew Strang, and Karon MacLean. 2015. Introducing the Cuddlebot: A Robot that Responds to Touch Gestures. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*. ACM, 295–295.
2. Lisa Feldman Barrett, Batja Mesquita, and Maria Gendron. 2011. Context in emotion perception. *Current Directions in Psychological Science* 20, 5 (2011), 286–290.
3. Susana Bloch, Madeleine Lemeignan, and Nancy Aguilera-T. 1991. Specific respiratory patterns distinguish among human basic emotions. *International Journal of Psychophysiology* 11, 2 (1991), 141–154.
4. Frans A Boiten, Nico H Frijda, and Cornelis JE Wientjes. 1994. Emotions and respiratory patterns: review and critical analysis. *International Journal of Psychophysiology* 17, 2 (1994), 103–128.
5. Paul Bucci, X. Laura Cang, Matthew Chun, David Marino, Oliver Schnieder, Hasti Seifi, and Karon E. MacLean. 2016. CuddleBits: an iterative prototyping platform for complex haptic display. *Eurohaptics Demonstration* (2016).
6. Laura Cang, Paul Bucci, and Karon E MacLean. 2015. CuddleBits: Friendly, Low-cost Furballs that Respond to Touch. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM, 365–366. DOI:http://dx.doi.org/10.1145/2818346.2823293
7. Jessica Q Dawson, Oliver S Schneider, Joel Ferstay, Dereck Toker, Juliette Link, Shathel Haddad, and Karon MacLean. 2013. It's alive!: exploring the design space of a gesturing phone. In *Proc of Graphics Interface 2013*. Canadian Information Processing Society, 205–212.
8. Celso M de Melo, Patrick Kenny, and Jonathan Gratch. 2010. Real-time expression of affect through respiration. *Computer Animation and Virtual Worlds* 21, 3-4 (2010), 225–234.
9. Lisa Feldman Barrett and James A Russell. 1998. Independence and bipolarity in the structure of current affect. *Journal of personality and social psychology* 74, 4 (1998), 967.
10. Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems* 42, 3 (2003), 143–166.
11. Jonas Forsslund, Michael Yip, and Eva-Lotta Sallnäs. 2015. WoodenHaptics: A Starting Kit for Crafting Force-Reflecting Spatial Haptic Devices. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, 133–140.
12. John Harris and Ehud Sharlin. 2011. Exploring the affect of abstract motion in social human-robot interaction. In *2011 Ro-Man*. IEEE, 441–448.
13. Fritz Heider and Marianne Simmel. 1944. An experimental study of apparent behavior. *The American Journal of Psychology* 57, 2 (1944), 243–259.
14. Tanja Hofer, Petra Hauf, and Gisa Aschersleben. 2007. Infants' perception of goal-directed actions on video. *British Journal of Developmental Psychology* 25, 3 (2007), 485–498.
15. Guy Hoffman and Wendy Ju. 2014. Designing robots with movement in mind. *Journal of Human-Robot Interaction* 3, 1 (2014), 89–122.
16. Youn-Kyung Lim, Erik Stolterman, and Josh Tenenberg. 2008. The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. *ACM TOCHI* 15, 2 (2008), 7.
17. C. Lithari, C. A. Frantzidis, C. Papadelis, Ana B. Vivas, M. A. Klados, C. Kourtidou-Papadeli, C. Pappas, A. A. Ioannides, and P. D. Bamidis. 2010. Are Females More Responsive to Emotional Stimuli? A Neurophysiological Study Across Arousal and Valence Dimensions. *Brain Topography* 23, 1 (2010), 27–40. DOI: http://dx.doi.org/10.1007/s10548-009-0130-5
18. Lars Mathiassen, Thomas Seewaldt, and Jan Stage. 1995. Prototyping and specifying: principles and practices of a mixed approach. *Scandinavian Journal of Information Systems* 7, 1 (1995), 4.
19. Camille Moussette. 2012. Simple haptics: Sketching perspectives for the design of haptic interactions. (2012).
20. Camille Moussette and Fabricio Dore. 2010. Sketching in Hardware and Building Interaction Design: tools, toolkits and an attitude for Interaction Designers. In *Design Research Society*.
21. S Oliver and M Karon. 2014. Haptic jazz: Collaborative touch with the haptic instrument. In *IEEE Haptics Symposium*.
22. Pierre Rainville, Antoine Bechara, Nasir Naqvi, and Antonio R Damasio. 2006. Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International journal of psychophysiology* 61, 1 (2006), 5–18.
23. Jussi Rantala, Katri Salminen, Roope Raisamo, and Veikko Surakka. 2013. Touch gestures in communicating emotional intention via vibrotactile stimulation. *Intl J Human-Computer Studies* 71, 6 (2013), 679–690.
24. Tiago Ribeiro and Ana Paiva. 2012. The illusion of robotic life: principles and practices of animation for robots. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 383–390.
25. James A Russell. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39, 6 (1980), 1161.
26. James A Russell, Anna Weiss, and Gerald A Mendelsohn. 1989. Affect grid: a single-item scale of pleasure and arousal. *J Personality & Social Psychology* 57, 3 (1989).

27. Jelle Saldien, Kristof Goris, Selma Yilmazyildiz, Werner Verhelst, and Dirk Lefeber. 2008. On the design of the huggable robot Probo. *Journal of Physical Agents* 2, 2 (2008), 3–12.
28. Oliver S Schneider and Karon E MacLean. 2016. Studying design process and example use with Macaron, a web-based vibrotactile effect editor. In *2016 IEEE Haptics Symposium (HAPTICS)*. IEEE, 52–58.
29. Oliver S Schneider, Hasti Seifi, Salma Kashani, Matthew Chun, and Karon E MacLean. 2016. HapTurk: Crowdsourcing Affective Ratings of Vibrotactile Icons. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3248–3260.
30. Yasaman S Sefidgar, Karon E MacLean, Steve Yohanan, HF Machiel Van der Loos, Elizabeth A Croft, and E Jane Garland. 2016. Design and Evaluation of a Touch-Centered Calming Interaction with a Social Robot. *IEEE Transactions on Affective Computing* 7, 2 (2016), 108–121.
31. Michael Shaver and Karon Maclean. 2003. *The Twiddler: A haptic teaching tool for low-cost communication and mechanical design*. Master's thesis.
32. Sotaro Shimada and Kazuo Hiraki. 2006. Infant's brain responses to live and televised action. *Neuroimage* 32, 2 (2006), 930–939.
33. Walter Dan Stiehl, Jun Ki Lee, Cynthia Breazeal, Marco Nalin, Angelica Morandi, and Alberto Sanna. 2009. The huggable: a platform for research in robotic companions for pediatric care. In *Proceedings of the 8th International Conference on interaction Design and Children*. ACM, 317–320.
34. Kazuyoshi Wada and Takanori Shibata. 2007. Living with seal robots—its sociopsychological and physiological influences on the elderly at a care house. *IEEE Transactions on Robotics* 23, 5 (2007), 972–980.
35. Kazuyoshi Wada, Takanori Shibata, Takashi Asada, and Toshimitsu Musha. 2007. Robot therapy for prevention of dementia at home. *Journal of Robotics and Mechatronics* 19, 6 (2007), 691.
36. David Watson, Lee A. Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54, 6 (1988), 1063–1070. DOI: http://dx.doi.org/10.1037/0022-3514.54.6.1063
37. Steve Yohanan and Karon MacLean. 2009. A tool to study affective touch. In *CHI'09 Extended Abstracts*. ACM, 4153–4158.
38. Steve Yohanan and Karon MacLean. 2011. Design and assessment of the haptic creature's affect display. In *HRI'11*. ACM, 473–480.
39. Steve Yohanan and Karon E MacLean. 2012. The role of affective touch in human-robot interaction: Human intent and expectations in touching the haptic creature. *Intl J Social Robotics* 4, 2 (2012), 163–180.