It's Alive! Exploring the Design Space of a Gesturing Phone

Jessica Q. Dawson*
University of British Columbia

Oliver S. Schneider*
University of British Columbia

Joel Ferstay*
University of British Columbia

Dereck Toker*
University of British Columbia

Juliette Link*
University of British Columbia

Shathel Haddad*
University of British Columbia

Karon MacLean*
University of British Columbia

ABSTRACT

Recent technical developments with flexible display materials have diversified the possible forms of near-future handheld devices. We envision smartphones that will deploy these materials for physical, device-originated gestural display as expressive channels for user communication. Over several iterations, we designed both human-actuated and mechanized prototypes that animate the standard block-like smartphone form-factor with evocative life-like gestures. We present three basic prototypes developed through an exploratory study, and a medium fidelity prototype developed in a second study, which enact a combination of visual and haptic gestural displays including *breathing, curling, crawling, ears,* and *vibration*. Through two evaluations we find that (a) users are receptive to the use of gestural displays to enrich their communications; and (b) smartphone-embodied gestural displays are capable of expressing both common notifications (*e.g.*, incoming calls) and emotional content through the dimensions of arousal and, to a small extent, valence. Finally, we conclude with several guidelines for the design of gestural mobile devices.

Keywords: Gestural display, body language, human-robot interaction, mobile computing, affective computing, ambient display, biological metaphor.

Index Terms: H5m [Information interfaces and presentation]: Miscellaneous.

1 INTRODUCTION

Compared to a decade ago, today's mobile devices are small and powerful. A combination of innovative interaction techniques and access to extraordinary amounts of information, both private and external, has increased the utility of mobile devices across many aspects of our lives. New developments in flexible display technologies promise to drive a similar transformation of the familiar rigid and rectangular form-factor of the typical smartphone.

The promise of portability, flexibility, and robustness make bendable graphical screens a target for industry innovation [1] and open up a myriad of interaction possibilities. Companies such as Samsung and Plastic Logic have used recent technologies, including flexible OLEDs (organic light-emitting diodes) and paper-like E-Ink, to develop displays that can be bent or flexed [2,3]. Existing prototypes of bendable mobile devices include Nokia's Kinetic Device [4] and the PaperPhone [5]. Fully functional flexible smartphones (manually flexed or squeezed) are expected to be commercially available by 2013 [6]. In turn, other technologies such as polymer membranes that contract or inflate/distend a surface [7], could lead to actuated (device-originated) motion for mobile devices.

* Email for all: {jqdawson, oschneid, joelaf, dtoker, jlink, shathel, maclean}@cs.ubc.ca

Figure 1: An early prototype (*Curl*) of a mobile device creating a biologically inspired life-like gesture.

These expressive capabilities also carry the potential to address the painful shortcomings in today's systems that stem from a lack of social intelligence: interruptions are either too intrusive or not salient enough. Regardless of the screen size or screen type, graphical notifications demand a user's attention, while audio notifications can be ineffective when quiet or muted, and socially inappropriate when loud. Meanwhile, haptic information transmission options are developing slowly, and current vibrotactile technologies are dimensionally narrow (duration, intensity, rhythm). There is clearly a need for alternative, more 'ambient' display modes, *i.e.*, interfaces that are natural, unobtrusive, and capable of operating in the background or periphery [8].

People are highly attuned to social and gestural cues [9], a facility long exploited for subtle, natural communication in animation, dance, and drama. We see physical life-like gestural display in smartphones as an added channel of communication that would leverage the unique capabilities of flexible, and eventually motion-capable, devices.

In this paper we focus on the rich *output* design space for a living metaphor enacted by smartphones to enhance channels of communication between the device and user. Over two studies, we created and evaluated four prototypes that explored a combination of life-like haptic and visual gestural displays including *breathing, curling, crawling, ears,* and *vibration*. Our first study sought to understand how users react to and interpret a focused set of physically rendered gestures displayed by a mobile device. Because we found that participants readily attributed emotional states to these various gestures, our second study focused on the design and evaluation of gestures intended to convey emotional range through the dimensions of valence and arousal. Our primary contributions are:

- Evidence that users are responsive to the living phone metaphor and feel that it could enrich their communication experience.
- Demonstration of smartphone-embodied gestural displays that are capable of conveying the emotional dimensions of arousal and, to a small extent, valence.
- Preliminary design guidance for the implementation of future gestural display behaviours in mobile devices.

2 RELATED WORK

Previous work has explored the use of expressive gestures and living metaphors for mobile displays, and the communication of emotion in mobile devices and social robots.

2.1 Expressiveness from Synthesized Gesture

Evoking human affective qualities by controlling the motion of non-human media has a long history enriched by graphical animation and film. Expressively animating simple, abstract shapes has a more focused history. Previous work demonstrates that humans anthropomorphize and attribute emotional characteristics when motion is applied to abstract shapes, including two-dimensional geometric objects like triangles, circles, and squares [10], or a physically animated simple wooden stick [11]. Indeed, gestures are a powerful medium of communication with the expressive capability of a complete verbal language [12].

2.2 Flexible displays and interaction for mobile devices

'Organic' and 'shape-changing' interfaces [13], where everyday objects often double as computation tools, have exploited flexible-displays to explore new handheld utilities and interactions. Several works have already investigated gestures like bending [5,13], shaking [15] and squeezing [16] as control inputs for mobile devices. However, expressive and informative ambient *display* is less studied, particularly with flexible display technology in mind. Hemmert et al. designed a series of ambient shape-changing displays that use the angling of rigid plates to deform a mobile device's body to communicate information for applications like navigation and user notifications [17,18]. To our knowledge, no work has yet examined the potential of embodied gestural communication in the context of mobile devices and flexible displays.

2.3 Living metaphors and emotion in mobile devices

Emotion is a natural fit for haptic and gestural displays, and a number of implementations have explored ways in which life-like signals can display notification information on mobile devices. Hemmert & Joost [19] designed an ambient vibrotactile display that simulated breathing and pulse in order to communicate two emotional states: 'calm' (normal pulse; no notifications), and 'excited' (high pulse; attention required!). Although users in their study disliked the constancy and intrusiveness of the pulsing implementation, they were generally receptive to the living metaphor. In terms of emotional expression, users interviewed by Heikkinen et al. felt that subtler emotional communication should be supported by haptic displays in mobile devices because such feelings are often difficult to express [20]. Studies have also explored alternative channels of communication in mobile devices. The eMoto service, although not haptic, explored the design of colors, shapes and animations inspired by body movements to express emotions in text messages [21]. As another example, Park et al. explored CheekTouch, by offering vibrotactile feedback of certain touches such as kissing or stroking through a user's cheek while talking on the phone [22]. The living metaphor has also been applied to mobile applications. For example, the anthropomorphic robot avatar CALLY [23] and the zoomorphic Cellular Squirrel [24] both use physical gestures for call management while considering social context. These animal-inspired forms capture a concrete biological metaphor, however, they lose the portability and discreetness of most mobile devices due to their size and shape.

2.4 Emotional displays in social robots

Abstract physical gestures for emotional displays have been studied in physical devices to help inform Human-Robot Interaction (HRI) and social robotics. Some studies have focused on creating emotion-conveying robots [*e.g.,* 25,26,27] as well as devices that add physical movements to products to enhance emotional expression [*e.g.,* 28,29]. Other research has explored how robot movements impact a user's perception of different emotions. Ju & Takayama explored how specific movements of an automatic door were interpreted as varying degrees of willingness to let people in [30]. Saerbeck & Bartneck found that the speed of the movements of a robotic cat and a Roomba vacuum could be used to predict users' perceived arousal of the robots, and that users had a similar interpretation for the movements of both robots, despite their quite different appearance [31]. The Haptic Creature [32], an actuated fur-covered robot, was designed to communicate purely through touch by employing motion, breathing, heartbeat, purring, and ear stiffening while sensing and then responding to user touch. A recent study with the Haptic Creature indicated that users could reliably agree on when the robot was content, but were less certain when it was upset or anxious [32]. The RobotPHONE teddy bear was similarly designed to express basic emotions using only arm and head motion; however, even with situational context emotion recognition rate was low (26% for 10 emotions) [33].

3 STUDY 1: EARLY DESIGNS OF LIVING GESTURES

Our initial exploration of the design space began with iterative *haptic sketching* [34], characterized by quick cycles of ideation, building, and evaluation of low-fidelity physical sketches. We then brought three of these low-fidelity physical prototypes to users in a qualitative study to gather feedback on the gestures and their affordances.

3.1 Exploring crawl, breathe, and curl gestures

We constrained our designs to the familiar rectangular shape of current smartphones to reduce the design space and focus on the effect of motion and materials. We deliberately avoided a fully anthropomorphic or zoomorphic look (*e.g.*, furry hand-held pet) that might excessively distort impressions of the object's purpose.

We began by brainstorming task examples and scenarios for ambient expressive gestures in mobile devices. Meant to drive exploration and imagination, these were not exhaustive but did encompass typical functions (*e.g.*, notification of incoming calls during an important meeting), and new possibilities (*e.g.*, communicate the urgency of a notification depending on the nature/sender of the call). Candidate gestures included the following behaviours: breathe, sit up, walk, look around, smile and wave, dance, shake, and pulsate.

In order to build our flexible physical prototypes, we set aside conventional rigid smartphone materials such as plastic and metal, and instead found inspiration in household materials like rubber, foam, and wood, which we could bend and move in life-like ways. We developed three prototypes (*Curl, Breathe*, and *Crawl*), based on the life-like movements that we found to be the most evocative, and which provided the richest exploration of materials and use cases. The three prototypes were each named for their primary gesture and were actuated by puppetry. We intentionally avoided manifesting all of our brainstormed gestures. After building these prototypes, the gestures covered the available materials, and the number of prototypes was feasible for a qualitative exploratory study.

All of the prototypes were approximately 6cm wide x 12 cm long (similar to the Apple iPhone 4), and varied in depth from 6 to 12mm. The *Curl* prototype (Figure 1) demonstrated a phone sitting up and looking at the user. Its hard, plastic body was segmented then bound with paper tape to support a smooth curl in one direction. Popsicle sticks were used for manual animation. The *Breathe* prototype (Figure 2-top) was constructed with grey closed-cell foam containing a balloon; a hard plastic back resisted the balloon's expansion. A puppeteer blowing in and sucking out of a breathing tube could animate the prototype to simulate breathing. The *Crawl* prototype (Figure 2-bottom) was a flattened, flexible shape with popsicle sticks at its ends for manipulation. Its freeform visual gestures included lateral movement along a table's edge (walking or crawling, inchworm-style) and looking around (by sitting or standing up).

3.2 Qualitative user study

To gauge acceptance of gestural display by a smartphone, explore gesture interpretations, and elicit feedback, we conducted a qualitative, exploratory user study with three previously acquainted pairs of participants (P1 to P6, 2 female) aged 23 to 49. Five of the participants had a computer science or HCI background. We referred to the prototypes by colour to reduce priming. Each of the three sessions took approximately 45 minutes and was video-recorded for later analysis.

3.2.1 Methods

The study was conducted in two semi-structured passes. In the first, the *Crawl*, *Curl*, and *Breathe* prototypes were presented to participants without context or reference to their intended movements. For each prototype, participants were first asked to discuss and describe their impressions. Next, the prototype was animated. Participants were then invited to touch and interact with the prototype, and describe their reactions. Experimenters asked open-ended questions without an intended type of feedback, and did not ask for emotional impressions. This process was then repeated for the next prototype, with presentation order counterbalanced across the three sessions. In the second pass, we revealed that the prototypes were intended to be smartphones. In the same order as the first pass, each prototype was again animated for the participants and they were asked to discuss and describe their impressions of the device.

Two researchers conducted independent video analyses of the video to generate feature sets of device affordances, compelling gestures and traits, and common interpretations or emotional significance of gestures. These three feature sets were used to develop *a priori* analysis criteria. As this was early in the design process, the video was analysed informally in three steps: 1) both researchers independently generated a vocabulary following the analysis criteria, then 2) met to consolidate their vocabularies, and finally 3) independently conducted a second confirmatory pass of the video, recording occurrences of each vocabulary component. This approach differs from formal coding with a single consolidation step rather than an agreement threshold, and a focus on emerging trends across participants.

Figure 2: Two early prototypes of living gestures, dubbed *Breathe* (top) and *Crawl* (bottom).

3.3 Results and Discussion

Analysis of the video-recordings revealed several important themes, around which participant responses are organized.

Animal and Living Metaphors: Both before and after being instructed to regard the prototypes as phones, participants consistently described the prototypes' motions with a wide range of primarily anthropomorphic metaphors.

All participants applied simple living metaphors to the prototypes. For instance, when facing upright *Curl* was seen as trying to be picked up or grabbing attention, and *Breathe* showed breathing or heart beating. Similarly, *Crawl* walked or crawled or wanted to be noticed. Such descriptions show that participants readily associated life-like behaviours with simple gestures.

There were also surprising and unexpected responses, often associated with more complex movements. These included metaphors based on movements and orientations that were tested during the evaluations – *e.g.,* "sea creature" (P2), "frightened" and "circus performer" (P3) for *Curl* when it was presented face down, "mouse" (P6) for Breathe in an instance where the air tube was prominent, "caterpillar" (P2) for *Crawl* – while others were based on the textures and feelings of the device – *e.g.*, "bones" and "scales" (P1). These responses demonstrate consistent anthropomorphic and zoomorphic interpretations, but also suggest that increased complexity led to broader variation in participants' specific interpretations of the gestures.

Orientation: Impressions of primary direction, pose, or orientation of each device dominated initial reactions. Participants routinely identified a 'head', and the direction it was facing. For example, interpretation of *Curl's* curling motion varied dramatically depending on whether the silver side of the curled end was facing the participant. Reactions included: "[It wants] your attention" (P6, when the device faced the participants) and: "Maybe it means it hates me, I've done something bad" (P4, when the device faced away). All six participants also felt that it was unnatural for *Curl's* perceived screen to face down on the table (*e.g.*, "he's uncomfortable") (P4). As further evidence for the importance of device orientation, most users interpreted *Breathe*'s movement as breathing or a heartbeat when the hard plastic side was downwards, but as "jumping" (P3) or "hopping" (P4) when the hard plastic side was upwards. The latter seemed more visually salient: "This is more attention grabbing" (P3). With the hard plastic side down, the first motion was more difficult to detect: "That breathing is practically invisible to you" (P2).

Urgency and Arousal: All participants related faster moving gestures as communicating both arousal and a sense of urgency. To illustrate, when either the breathing motion was sped up, or the curl-up motion was performed more quickly, participants always identified this as an increased level of excitement, which in turn conveyed a greater sense of urgency.

3.4 Summary

These results offer a first glimpse of insights that could guide further design of smartphone-embodied gestural displays, and are elaborated on in Section 5. Simplicity, orientation, salience, and speed of motion all impacted the consistency of the gesture interpretations related by participants. Some interpretations, such as a face-up *Curl* wanting attention or to be picked up, suggest that simple device-originated gestures may map effectively to use cases that describe incoming notifications. Most importantly, we found that emotional and metaphoric attribution came easily to participants, who readily volunteered many living metaphors and ascribed emotions such as hate and excitement to the gestures.

4 Study 2: Designing gestural displays for emotion

Related research has shown that humans readily attribute emotion to physical movements performed by inanimate objects [11,12] and findings from our first study suggest that this may also be the case for mobile devices. With this knowledge, we decided to explore the possibility of using gestural display to enrich the communication channel between a user and their smartphone by adding emotional content. Building upon the provocative emotional connections that participants made in our first study, we felt that emotional context was a natural use case for a more focused exploration of the design space of physical gestures for smartphone displays.

We chose augmentation of text messages as an ideal starting point since voice, text, and occasionally images (the current palette for exchanges between mobile users) do not support all of the nuances of physically present person-to-person interactions. For example, verbal prosody, facial expression, and body language all add important emotional context that enriches communication and without them, a text message like "I'm fine" is ambiguous or flat. The use of emoticons (e.g. "I'm fine :)" and "I'm fine :p") are one widespread partial solution which many use on a daily basis [35], but they are limited. Video chat has addressed many other constraints to expressiveness, but is unsuited for notifications and asynchronous communication.

4.1 Selecting gestures for emotional display

To address the lack of consistency in participant interpretations seen in our first study, we strove for a small set of gestures that could be intuitively and consistently mapped to specific emotions. The emergence of urgency/arousal as a major theme in the first study, and the variety of positive and negative interpretations of gestures suggested existing two-dimensional models of affect to be suitable [36]. The process used by Yohanan and Maclean in developing the affective display for the Haptic Creature [37]

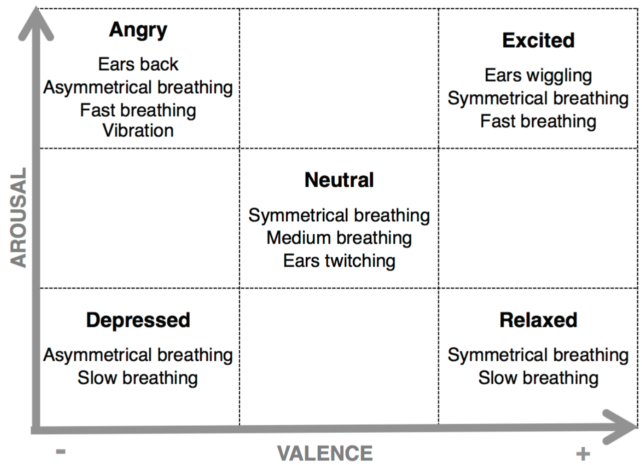

Figure 3: DEVA expressions used in 2nd evaluation, with *ears, breathing*, and *vibration* modalities, shown on Russell's two-dimensional model of emotion [36] as adapted by [37].

inspired our gesture design and evaluation, as well as our use of a version of Russell's two-dimensional model of emotion [36], which combines two scales, valence and arousal.

Valence spans negative (sad) to neutral to positive (happy) along the horizontal axis. Arousal, a measure of excitement or activity, spans from low to medium to high along the vertical axis. Although higher dimensional models are required to capture more subtle nuances in emotion, this two-dimensional model is widely used in the psychological community [38] and provided enough resolution for exploration.

Emotions can be placed along these scales by their arousal and valence levels (Figure 3), similar to Lang's Self-Assessment Manikin (SAM) scale [39]. This representation supports a systematic, combinatorial approach to assembling the display of specific emotions: first target a point in the valence/arousal space based on the expressiveness found in individual display elements, then verify or iterate the combination for synergistic shifts.

Arousal: We chose breathing to express a range in arousal based on previous work [19] and the readiness with which participants connected breathing speed to urgency in Study 1.

Valence: As our previous study did not specifically address valence, we chose to explore an ear-flexing gesture based on the finding of [37] that participants connect robot ears intended to express arousal with valence instead. We also included vibration, as most mobile devices have this capability, and shaking and shivering are another effective metaphor for communicating negative valence [37].

Through iterative cycles of haptic sketching [34] we created a number of physical sketches (*breathing*, *ears*, and *vibration*) which we actuated as puppets (excluding *vibration*). Our initial puppet sketches of *ears* were based on observations of cat ears. Small movement irregularities due to human actuation were ideal for exploration.

4.2 Medium fidelity prototype design: DEVA

Our haptic sketching culminated in a final prototype: *DEVA* (Device Expressing Valence and Arousal, see Figures 4 & 5). DEVA is made of a dense black foam and measures 7cm wide x 13cm long x 3cm deep, similar to existing phones. We mechanized all gestures to support consistency and coordination through the use of two Arduino Uno microcontrollers [40] which controlled the embedded actuators.

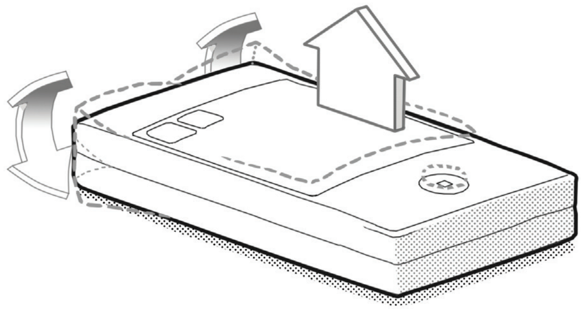

Figure 4: A sketch of the final prototype, DEVA, illustrating the location and range of motion for the *ears* and *breathing*

Figure 5: The complete DEVA prototype

The gestures *ears* and *breathing* were both implemented with servos, while *vibration* was achieved using a 1cm in diameter buzzer motor placed inside and midway along the body's axes. Figure 4 illustrates the location and range of motion of the servos controlling the *ears* and *breathing*. The *ears* were created with two servos by sewing the servo housing and levers into cavities carved into the corners of the foam body, and were not visually apparent when at rest. The servo levers, grounded against the foam, pushed the corners of the prototype forward or backward at varying speeds like twitching ears. *Breathing* was implemented with another small servo hidden in the lower middle of the body. Placed against the rigid back plate of the device, it pushed up a plastic hemisphere to render inhalation and exhalation.

The prototype needed to be plausible as a smartphone. For a more phone-like (*i.e.*, firm and solid) feel, we enclosed the back and sides in a firm plastic case, with a sheet of aluminium in the back for weight and rigidity. We printed an image of an iPhone onto cotton fabric and affixed it to DEVA's face (Figure 5).

We devised five emotional expressions – *angry, depressed, neutral, excited*, and *relaxed* – to cover the four corners of the two-dimensional model as well as the neutral middle state (Figure 3). This selection of expressions ensured that each of the three levels of arousal (low, medium and high) and the three levels of valence (negative, neutral, positive) were represented at least once. Each expression employed some combination of *ears*, *breathing* and/or *vibration*. DEVA varied *breathing* rate and symmetry for each expression, as in [37]. Faster or slower breathing rate expressed high and low arousal, respectively. Symmetrical, even breaths were used for positive valence, and asymmetrical breaths – a long inhale, with a shorter exhale, as in a sigh or huff – were used for negative valence. In pilots, *ears* activity was associated with higher arousal, so they are active only for high arousal expressions; *vibration* was perceived as jarring, and thus reserved for high arousal, negative valence.

4.3 Evaluating gestures for displaying emotion

We conducted a small user study to examine how well the life-like gestures could express emotion within the valence/arousal framework of DEVA. We presented DEVA's five programmed expressions to users, and measured the extent to which participant ratings of arousal and valence for each expression matched up with the targeted ranges. The study was conducted with 10 participants (P1 to P10, 4 female) aged 24 to 67. Three of the participants had a computer science or HCI background.

4.3.1 Methods

Each study session took approximately 20 minutes. First, participants were introduced to the prototype as a phone that displays the emotion of an incoming text messages through *ears* and *breathing* gestures. They were asked to hold the prototype in one or both hands and shown each of the five expressions.

Next, participants were again shown each of the five expressions and asked to report the arousal and valence perceived in each, using nine-point versions of the SAM scales for valence and arousal [39]. Presentation order was counterbalanced with a partial Latin square. The study concluded with an interview where the participant was invited to describe his or her reaction to the gestures, the concept of smart-phone embodied gestural displays, and the prototype. Participants were also asked to comment on the realism of the use case of showing text message emotion, and to suggest other scenarios for emotional display through their smartphone. As these sessions were focused on obtaining participant ratings of the emotional expressions and not on exploring interactions, we did not employ video.

4.4 Results and Discussion

The means and standard deviations of the SAM scale scores for each expression are shown in Figure 6. To analyse the scores, we ran one-way repeated measures ANOVA for the expression factor for both arousal and valence. In order to ensure that our data was valid for standard parametric analyses, the Aligned Rank Transformation was applied to the valence and arousal ratings to ensure consistency over monotone transformations [41]. Finally, we also plotted the perceived arousal and valence scores (Figure 7) to help illustrate any patterns and support exploration of unexpected results.

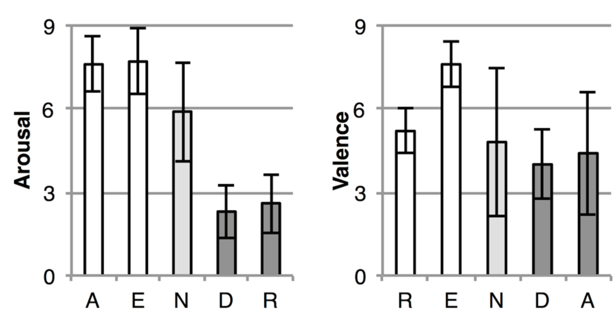

Figure 6: Average arousal and valence scores for the five expressions, labelled by the first letter of the expression name and grouped according to the targeted level. Left: high (white), medium (light), low (dark) arousal expressions. Right: positive (white), neutral (light), negative (dark) valence expressions.

Table 1: Perceived arousal homogeneous subsets

Condition	N	Subset for Alpha = .05		
		1	2	3
Depressed	10	2.30		
Relaxed		2.60		
Neutral			5.90	
Angry				7.60
Excited				7.70
Sig.		.982	1.000	1.000

Table 2: Perceived valence homogeneous subsets

Condition	N	Subset for Alpha = .05	
		1	2
Depressed	10	4.00	
Angry		4.40	
Neutral		4.80	
Relaxed		5.20	
Excited			7.60
Sig.		.533	1.000

4.4.1 SAM scale scores

DEVA successfully communicated levels of arousal. The one-way repeated measures ANOVA yielded a statistically significant difference among the 5 expressions for arousal $F(4,36) = 56.96$, $p < .001$, $\eta p^2 = 0.86$. Multiple comparisons via Tukey's HSD computed three homogeneous subsets (Table 1). These results match up with our targeted outcome: the first subset contained the low arousal expressions (*depressed*, *relaxed*), the second subset contained the medium arousal expression (*neutral*), *and* the third subset contained the high arousal expressions (*angry*, *excited*). The subsets are illustrated in Figure 6-left by the step pattern from the high arousal expressions down to the low arousal expressions.

Figure 7 shows how the low and high arousal expression responses cluster at the low and high ends of the scale. *Neutral* largely spans the centre of the scale, but was somewhat less consistently perceived than the other four expressions. Breathing seems to have been effective for communicating arousal as expected, and from participant responses we suspect that the presence of the twitching *ears* likely also emphasized the high arousal expressions.

DEVA successfully expressed higher valence for excited. The one-way repeated measures ANOVA yielded a statistically significant difference among the 5 expressions for valence $F(4,36) = 7.66$, $p < .001$, $\eta p^2 = 0.46$. Multiple comparisons via Tukey's HSD computed two homogeneous subsets (Table 2): the second subset contained *excited,* while the first subset contained the remaining four expressions. We had expected three subsets, one for each level of valence (negative, neutral, positive).

As Figure 6-right shows, *excited* was the only expression successful in communicating a difference in valence. Figure 7 helps to shed light on these results. The *excited* scores cluster as expected at the right (positive) side of the scale. *Angry* and *depressed* were somewhat consistently interpreted, but the scores largely cluster on the boundary between negative and neutral valence. *Relaxed* was very consistently perceived, but the scores cluster in the central (neutral) region, rather than to the right (positive) as expected. The *neutral* expression was the most inconsistent, with a large variance in scores.

The different *ear* movements may partially explain why DEVA was not successful in communicating the expected range in valence: the *ears* in the *excited* expression were perceived as energetic by participants, while a combination of backward folding *ears* and *vibration* were largely perceived as indicative of tension or stress for *angry* (except for two participants who again perceived this as energetic). The asymmetric breathing appears to have been somewhat effective in communicating *depressed* and *angry*, but this effect alone was not strong enough to ensure either expression was definitively negative to everyone. *Depressed and relaxed* were less clearly differentiated than *angry* and *excited,* which might be attributable to the expressions' over-reliance on variation in *breathing* symmetry, and that neither used *ears* or *vibration*. Finally, *neutral* used the middle of *ears*' and *breathing*'s speed and range of motion rather than the extremes - the *neutral* expression may be more effective without *ears,* using just symmetrical *breathing* at a slightly faster rate than what was used for the original *relaxed* expression.

4.4.2 Follow-up Interview

The metaphors of *breathing* and *ears* were overall clear to participants, and were largely seen as believable representations for how a living phone would enact such gestures. P9 said, "*It's pretty realistic for a phone that breathes. It even has a little belly. When we started, it seemed lumpy, but as I'm interacting it seems more human. I'm putting human and animal features on it*." However, as in our first study, some participants also offered other interpretations for the gestures, which were often influenced by exactly how the prototype was held: P7 noted that sometimes the *breathing* felt more like a heartbeat, while P8 commented that if the prototype was held upside down, the *ears* became more like feet. Finally, although participants were able to both see and feel the gestures in our study, P6 and P8 noted that it would be important for them to be able to just feel the gestures (*e.g.*, in a pocket).

A phone capable of expressing emotion was palatable to eight of the ten participants. P1 specified that he was not comfortable with showing the emotion of a text because he did not want others to have control over what his phone does, but was interested in using the feature in other contexts such as assigning "emotional ringtones" to callers (*e.g.*, a "happy" ringtone that indicates a call from a friend). Other suggestions for applications offered by participants included signalling urgency for messages (*e.g.*, based on relationship with the sender) or time sensitive notifications (*e.g.*, alarms) and device status (*e.g.* battery life), and improving notifications for other applications by adding emotional context (*e.g.*, stocks, map localization).

Two participants said they were unlikely to use a gestural mobile device: P5 was not fond of notifications in general because he found all notifications disruptive, and P9 thought such a device might prove problematic to configure. P4 liked the device, but expressed concern about constant movement, preferring that it remain stationary except when communicating.

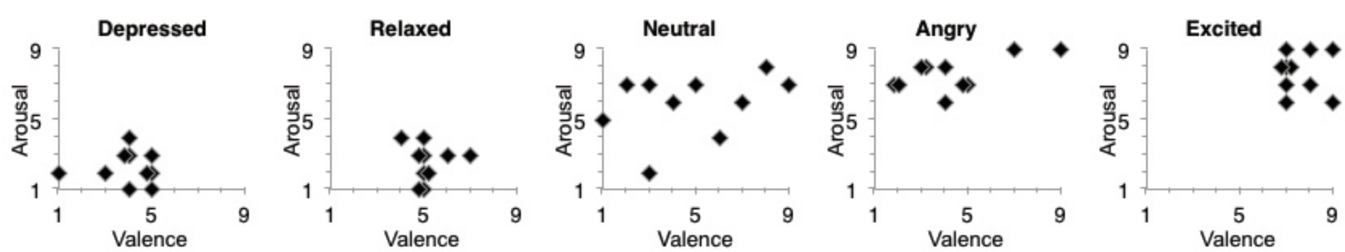

Figure 7: Scatterplots of perceived arousal and valence scores for each of the five expressions.

4.5 Summary

While DEVA communicated valence effectively only for the *excited* expression, the prototype successfully communicated three levels of arousal as expected. These results are heartening, and we believe that we can improve DEVA's ability to communicate valence through further refinement. The results also add to our understanding of how breathing speed and symmetry may map to arousal and valence, and provide some evidence that ears may help to communicate valence as intended. Based on these findings, we should be able to design more effective versions of *relaxed* and *neutral*. This could include adding *ears* to all of the expressions, *e.g.*, ears moving backward slowly for *depressed* and slowly forward for *relaxed,* and additional or different gestures to add clarity. These findings build on previous work mapping simple biological gestures to arousal and valence to communicate consistent emotions, and although preliminary, suggest that mobile devices may be a suitable platform for gestural emotional expression.

Finally, as in the first study, our interviews with participants suggest that users readily anthropomorphized the gestures and reinforce our earlier observation that explicit animal features may not be necessary to achieve such an effect. The participants largely found the movement and expressions of DEVA organic. Most of the participants expressed interest in a gestural phone, and even envisioned applications for emotional notifications beyond sending the emotions of text messages.

5 IMPLICATIONS FOR DESIGN

From our design exercises and two studies, we have compiled preliminary implications for the design of smartphone-embodied gestural displays.

Simple gestures go a long way. Our simplest gestures proved the most powerful and consistent. Simple gestures also seem to be more easily mapped to common smartphone functions, such as *Curl* wanting to be "picked up" for an incoming call or alert, or the speed of *Breathe* to communicate a notification's urgency.

We observed that each of our gestures naturally give a sense of urgency or arousal based on rate or speed. Participants connected device anxiety to breathing speed, as shown in [19], but also to the speed of the *Curl* and *Crawl* gestures. Building on previous work that correlates speed of movements with perceived arousal [31], this suggests that many gestural displays in addition to breathing or pulsing could communicate a notification's urgency through changes in rate. Because simple gestures such as breathing or sitting-up are quite natural, changes in the rate or speed can be intuitively mapped to personal experience. Designers should seek to understand and leverage common interpretations of gestures, and assign them to notifications or novel functions accordingly. More complex or asymmetrical gestures are less likely to have a common, natural mapping, and may elicit a wider variety of interpretations.

Orientation can impact the meaning of gestures. Many gesture interpretations changed with orientation. When *Curl* faced users, they said it was trying to get their attention; when it faced away it was "hiding something", and when face down it looked "uncomfortable." *Breathe* only showed its intended gesture when face up, but "hopped" or "jumped" when facing down. Similarly, one participant perceived DEVA's *ears* as legs when the prototype was upside down. We also found that asymmetric gestures like breathing are often easier to notice (visually and/or haptically) in one orientation rather than another. Designers should bear in mind the contributions of both orientation and user viewpoint to perceived meaning and salience of physical gestures.

Gestures may be perceived differently when 'seen,' 'felt', or both. For example, we observed that the *Curl* prototype lost some expressiveness if it was not seen while moving; when it was held, the complete range of motion was perceptible and less likely to go unnoticed. The same was observed with DEVA's *breathing* gesture, which one participant found more faint and heartbeat-like when held certain ways. Designers should therefore consider both look and feel when designing biological gestures, and ensure that these work as expected alone and in combination.

Transient gestures are preferred, but may be easily missed. Our evaluations used transient gestures: they were shown to the user within a short time frame while actively interacting with the prototype, and then the gesture would end. In real-life, mobile devices are often in pockets or bags, and are not always within visual or physical reach: users could therefore miss important cues from transient gestures such as speed, especially if the device returns to its original position at the end. However, feedback from some users in the second study suggests they would be annoyed if their device used continuous gestures and was always moving. Designers should consider these trade-offs carefully when deciding on triggers, transience, and frequency for gestures.

5.1 Reflections on the design process

An arousal/valence framework helps to map gestures to emotions. Our process benefitted from the use of an established framework [18, 23] for both the design and evaluation of our emotional gestures; this structure gave us a broader frame of reference with which to understand our designs.

Providing context is important when soliciting feedback on gestures. Participants tended to view our early prototypes as little creatures or other anthropomorphic devices, but by making DEVA look and feel more like a mobile device, participants were more able to provide imagined situations using the device.

Puppetry and haptic sketching are powerful design tools. The process of doing multiple iterations and creating quick physical sketches based on the qualities of available materials inspired us to explore gestures we may not have otherwise.

6 CONCLUSION AND FUTURE WORK

In this work, we sketched a series of prototypes that explore the design of a smartphone that communicates though life-like haptic and visual gestural displays including *breathing, curling, crawling, ears*, and *vibration*. Through an exploratory evaluation we observed a readiness in participants to anthropomorphize our early prototypes and to assign emotional intentions to the gestures. We leveraged this behaviour to design and evaluate combinations of *breathing, ears,* and *vibration* gestural displays, to effectively communicate the emotional dimensions of arousal and, to a small extent, valence. From our overall findings, we assembled a set of guidelines that provide insight for including gestural display into the mobile device design-space. While not all of the gestures and expressions in the two studies were always interpreted according to our expectations, our evaluations suggest that gestural displays have great potential to enrich our device-mediated communications through expressiveness and emotion.

Although the selected gestures tackle only a fraction of the gestural design space afforded by flexible displays, our results provide a foundation for the initial exploration of smartphone-embodied gestural displays. There are many ways in which a breathing gesture might be implemented, and there are likely many possible gestures that could be tailored to specific functions. It will also be crucial to understand how gestural displays can work in combination with existing output displays like sound and ringtones, as well as novel non-gestural functions, like changing colour or temperature. Ideally, we would like to move beyond household materials to work with existing flexible displays to understand the constraints and opportunities inherent in the real materials of the next generation of smartphones.

Future work should follow up on the exploratory nature of our user studies with more rigorous evaluations of the effectiveness of individual and combined gestures for expressing emotions. This would include investigating if the display can be expanded to express a broader range of emotions with more nuanced differences, and determining if DEVA's communication of valence can be improved simply by changing the programming of the existing gestures. We will also need to explore the semantics of these gestures to determine which types of notifications, tasks, and contexts are best suited for which gestures.

Finally, while this work focused on channels for device output and notifications, it is important to understand interactivity as a whole. What would it mean if a person 'squashed' the sitting-up phone, or tickled it, or twisted it? Future studies should continue to examine how flexible displays can be best leveraged for input through gestures like squeezing and bending, and how input and output gestures can be best combined to support fluid interactions.

7 ACKNOWLEDGMENTS

We thank Alexander Boswell for creating the sketch of DEVA (Figure 4). This work was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), and Nokia Inc.

REFERENCES

[1] M.-C. Choi, Y. Kim, and C.-S. Ha, "Polymers for flexible displays: From material selection to device applications," *Progress in Polymer Science*, vol. 33, no. 6, pp. 581–630, Jun. 2008.

[2] P. Ganapati, "Samsung shows flexible display for cellphones," 2008.

[3] "Plastic Logic Technology," http://www.plasticlogic.com, 2012.

[4] T. Davies, "Nokia Kinetic bendy phone is the next big thing," *http://conversations.nokia.com*, 2011.

[5] B. Lahey, A. Girouard, W. Burleson, and R. Vertegaal, "PaperPhone: understanding the use of bend gestures in mobile devices with flexible electronic paper displays," in *Proc. CHI '11*, 2011, pp. 1303–1312.

[6] K. Moskvitch, "Bend me, shape me: Flexible phones 'out by 2013'," *www.bbc.co.uk*, 2012.

[7] S. Kim, H. Kim, B. Lee, T.-J. Nam, and W. Lee, "Inflatable mouse: Volume-adjustable Mouse with Air- pressure-sensitive Input and Haptic Feedback.," in *CHI'08*, 2008, pp. 211–214.

[8] K. E. MacLean, "Putting Haptics into the Ambience," *IEEE Trans. Haptics*, vol. 2, no. 3, pp. 123–135, Jul. 2009.

[9] C. Ennis, R. McDonnell, and C. O'Sullivan, "Seeing is believing: body motion dominates in multisensory conversations," *ACM ToG*, vol. 29, no. 4, pp. 91:1–91:9, Jul. 2010.

[10] F. Heider and M. Simmel, "An Experimental Study of Apparent Behavior," *The American Journal of Psychology*, vol. 57, no. 2, pp. 243–259, 1944.

[11] J. Harris and E. Sharlin, "Exploring emotive actuation and its role in human-robot interaction," in *Proc. HRI '10*, 2010, pp. 95–96.

[12] A. Kendon, *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004, p. 400.

[13] M. K. Rasmussen, E. W. Pedersen, M. G. Petersen, and K. Hornbæk, "Shape-changing interfaces," in *Proc. CHI '12*, 2012, p. 735.

[14] J. Kildal, S. Paasovaara, and V. Aaltonen, "Kinetic device: designing interactions with a deformable mobile interface.," in *CHI EA '12:*, 2012, pp. 1871–1876.

[15] R. Strong and B. Gaver, "Feather, Scent, and Shaker: Supporting Simple Intimacy," in *CSCW '96*, 1996, pp. 29–30.

[16] B. L. Harrison, K. P. Fishkin, A. Gujar, C. Mochon, and R. Want, "Squeeze me, hold me, tilt me! An exploration of manipulative user interfaces," in *CHI '98*, 1998, pp. 17–24.

[17] F. Hemmert, S. Hamann, M. Löwe, J. Zeipelt, and G. Joost, "Shape-Changing Mobiles: Tapering in One-Dimensional Deformational Displays in Mobile Phones," in *Proc. TEI'10*, 2010, pp. 249–252.

[18] F. Hemmert, S. Hamann, M. Löwe, J. Zeipelt, and G. Joost, "Shape-changing mobiles: tapering in two- dimensional deformational displays in mobile phones.," in *CHI EA'10*, 2010, pp. 3075–2079.

[19] F. Hemmert and G. Joost, "Ambient Life: Calm and Excited Pulsation as a Means of Life-like Permanent Tactile Status Display in Mobile Phones," in *Proc. Design & Emotion '08*, 2008.

[20] J. Heikkinen, T. Olsson, and K. Väänänen-Vainio-Mattila, "Expectations for user experience in haptic communication with mobile devices," in *Proc. MobileHCI '09*, 2009, pp. 28:1–28:10.

[21] A. Staahl, P. Sundström, and K. Höök, "A foundation for emotional expressivity," in *Proc. DUX '05*, 2005.

[22] Y.-W. Park, C.-Y. Lim, and T.-J. Nam, "CheekTouch," in *Proc. CHI EA '10*, 2010, p. 3241.

[23] J.-D. Yim and C. D. Shaw, "Designing CALLY,: a cell-phone robot," in *CHI EA '09*, 2009, pp. 2659–2662.

[24] S. Marti and C. Schmandt, "Physical Embodiments for Mobile Communication Agents," in *Proc. UIST '05*, 2005, pp. 231–240.

[25] C. Breazeal, "Emotion and sociable humanoid robots," *International Journal of Human-Computer Studies*, vol. 59, no. 1–2, pp. 119–155, Jul. 2003.

[26] H. Lim, A. Ishii, and A. Takanishi, "Emotion-based biped walking," *Robotica*, vol. 22, no. 5, pp. 577–586, Oct. 2004.

[27] H. Mizoguchi, T. Sato, K. Takagi, M. Nakao, and Y. Hatamura, "Realization of Expressive Mobile Robot," in *Proc. Robotics and Automation*, 1997, vol. 1, pp. 581–586.

[28] J. Lee and T. Nam, "Augmenting Emotional Interaction Through Physical Movement.," in *Proc. UIST '06*, 2006.

[29] H. Kazuhiko, "Thanks Tails." 1996.

[30] W. Ju and L. Takayama, "Approachability: How people interpret automatic door movement as gesture.," *International Journal of Design*, vol. 3, no. 2, pp. 1–10, 2009.

[31] M. Saerbeck and C. Bartneck, "Perception of affect elicited by robot motion.," in *Proc. Human-Robot '10*, 2010, pp. 53–60.

[32] S. Yohanan and K. E. MacLean, "The Role of Affective Touch in Human-Robot Interaction: Human Intent and Expectations in Touching the Haptic Creature," *International Journal of Social Robotics*, vol. 4, no. 2, pp. 163–180, Dec. 2011.

[33] J. Li and M. Chignell, "Communication of Emotion in Social Robots through Simple Head and Arm Movements," *International Journal of Social Robotics*, vol. 3, no. 2, pp. 125–142, Sep. 2011.

[34] C. Moussette, "Learn to make, make to learn: Reflections from Sketching Haptics Workshops.," in *Proc. DeSForM '12*, 2012.

[35] R. R. Provine, R. J. Spencer, and D. L. Mandell, "Emotional Expression Online: Emoticons Punctuate Website Text Messages," *Journal of Language and Social Psychology*, vol. 26, no. 3, pp. 299–307, Sep. 2007.

[36] J. A. Russell, "A circumplex model of affect.," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[37] S. Yohanan and K. E. MacLean, "Design and assessment of the haptic creature's affect display," in *Proc. HRI '11*, 2011, pp. 473–480.

[38] J. R. J. Fontaine, K. R. Scherer, E. B. Roesch, and P. C. Ellswort, "The World of Emotions Is Not Two-Dimensional," *Psychological Science*, vol. 18, no. 12, pp. 1050–1057, 2007.

[39] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8.," Gainesville, FL., 2008.

[40] "Arduino," www.arduino.cc, 2012.

[41] J. Wobbrock, L. Findlater, D. Gergle, and J. Higgins, "The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures.," in *Proc. CHI '11*, 2011.