Some Like it Hot? Thermal Feedback for Mobile Devices

Graham Wilson, Martin Halvey, Stephen A. Brewster
School of Computing Science, University of Glasgow,
Glasgow, United Kingdom G12 8QQ
gawilson@dcs.gla.ac.uk,{first.last}@glasgow.ac.uk

Stephen A. Hughes
SAMH Engineering, 1 Leopardstown
Drive, Blackrock, Dublin Ireland
stephenahughes@gmail.com

ABSTRACT
Thermal stimulation is a rich, emotive and salient feedback channel that is well suited to HCI, but one that is yet to be fully investigated. Thermal feedback may be suited to environments that are too loud for audio or too bumpy for vibrotactile feedback. This paper presents two studies into how well users could detect hot and cold stimuli presented to the fingertips, the palm, the dorsal surface of the forearm and the dorsal surface of the upper arm. Evaluations were carried out in static and mobile settings. Results showed that the palm is most sensitive, cold is more perceivable and comfortable than warm and that stronger and faster-changing stimuli are more detectable but less comfortable. Guidelines for the design of thermal feedback are outlined, with attention paid to perceptual and hedonic factors.

Author Keywords
Thermal feedback, mobile interaction, non-visual feedback.

ACM Classification Keywords
H5.2. User Interfaces: Haptic IO

General Terms
Design, Human Factors

INTRODUCTION
Modern mobile devices are almost uniformly moving towards interfaces based on physical interactions such as multitouch and 3D spatial input through gestures. Although the use of haptic feedback has received a lot of attention, this has mostly been through the vibrotactile actuators built into most phones. There has been very little work in HCI on other aspects of the touch sense, in particular the thermal aspects, which we focus on in this paper.

There are many potential uses for thermal feedback; the following usage scenarios provide some potential examples:

"Jo is travelling to a meeting on the train. She is carrying a mobile device with personalised travel apps. Before her train arrives she has a drink near the station. After some time, her device begins to slowly warm up indicating that her train is nearing the station. She finishes up and heads for the train. As she boards the train, thermal feedback provides information regarding proximity to her seat; the device gets warmer as she moves towards the seat and colder as she moves away. This non-visual feedback would mean that she doesn't need her ticket to hand to check seat numbers as she walks down the train."

"Ted is in the library. He receives a text message and his phone vibrates in his pocket, but the rumbling of the phone off the seat disturbs people close by. Ted takes out his phone, slightly embarrassed he sets his phone to give him just thermal notifications. Now all messages, etc. will be private only to Ted. As one particular message is received his phone cools, indicating that an important message has been received. He leaves the library to check his message"

These examples illustrate the numerous potential uses for thermal feedback. It can act as an alternative non-visual notification channel for situations that are too bumpy or noisy for vibrotactile and audio feedback. It can augment both visual and non-visual feedback to add an extra richness to the interaction experience. In addition, thermal output is also entirely private; in contrast, vibrotactile feedback can sometimes still be heard or felt by others.

Much work has examined the underlying characteristics of human thermal perception and this is not the focus of this paper. Fundamental HCI research needs to be conducted to properly judge the merits of thermal feedback as a practical display method. It is necessary to understand the stimulus characteristics (e.g. intensity and rate of change) that are most effective for arousing sensation in the user. We focus on designing structured feedback for HCI that produces detectable stimuli using hardware designed for use with mobile devices. Two studies were conducted to test how well users could detect warm and cool changes in thermal sensation presented to the hand and arm; one study focused on sitting and the other walking around an indoor environment. Note that in this paper we use the terms 'warm' and 'cool/cold' (e.g. in 'warmth/cold perception') to refer to *warming* or *cooling*: increases or decreases in temperature.

BACKGROUND
Temperature Perception
The human skin rests in a relatively small 'neutral' homeostatic thermal state, ranging from around 28°C up to a maximum of 40°C when in all but the most extreme environmental conditions [12, 22]. The size of this neutral zone is relatively constant across individuals at around 6-8°C, but due to individual differences in thermal sensitivity, the relative position of each individual's neutral zone varies

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

(e.g. 30-36°C, 28-34°C, etc.). Within the neutral zone there is no discernable thermal sensation [12], i.e. there is no sensation of warmth or cold, and *adaptation* (where the sensation of thermal neutrality returns after heating or cooling the skin to a different temperature) only occurs within this range [22]. Outside of this range a constant sensation of warmth (above) or cold (below) is perceived [15]. Jones & Berris [12] suggested that slow changes in thermal stimulation within the neutral zone are difficult to detect. Therefore, for thermal stimuli to be reliably detected the skin must be presented with temperatures out-with this zone or changes in thermal stimulation must be rapid. There are significantly (approximately 30x) more cold than warm receptors [22] and the speed of cold perception is faster than warmth perception. Kenshalo [13] suggests that cold perception has a more immediate onset whereas warm sensations grow slowly before 'blooming'.

Thermal *just noticeable differences* (JNDs) refer to the minimum amount of change (warming or cooling) from current skin temperature that is required for that change to be detected, and are similar to tests of just noticeable differences in other perceptual modalities. For example, a JND of 1.5°C means any change of less than 1.5°C cannot be felt and, in general, the smaller the JND, the more salient the stimulus. JNDs are measured from a set baseline temperature and are inextricably linked to both this starting temperature and the rate of change (ROC) of the stimulus. At low rates (up to 3°C/sec) the size of both warm and cold JNDs decrease as ROC increases, with the most dramatic decrease occurring from ~0.01°C/sec to 0.3°C/sec. Above 3°C/sec (up to ~7°C/sec) JNDs then begin to increase again, with this being attributed to the conduction velocities of thermal receptors [2, 20] as well as reaction and cognition time [8]. This suggests that stimuli become more salient as ROC increases; however, above a set speed, even if salience increases, the ROC 'overtakes' reaction time so that further increases in stimulation have occurred by the time the participant could react. As the skin adapts to the warm or cool extremes of the neutral zone, warm and cold JNDs, respectively, decrease and decrease more as the stimulus intensity approaches the heat/cold pain thresholds (~45°C and 11-15°C respectively [11, 22]). In other words, we become more sensitive to thermal changes if they move the skin closer to pain thresholds. Conversely, warm and cold JNDs increase as the skin is cooled and warmed respectively. From this, and other evidence, it is clear that the thermal sense is more sensitive to *changes* in temperature, rather than absolute temperature itself.

Thermal sensitivity is not uniform across the body; there are marked variations between different locations as well as between different skin types. Glabrous skin (hairless skin as found on the fingertips or palm) is generally less sensitive to changes in thermal stimulation than non-glabrous, hairy, skin, with JNDs being generally larger on glabrous skin due to skin thickness [8, 20]. The thenar eminence (the bulbous region of the palm adjoining the thumb) has higher sensitivity than the rest of the palm [7], but is still not as sensitive as non-glabrous skin on the hand [8]. In general, thermal sensitivity is best on the head and trunk but worse towards the extremities [2, 7].

The JNDs mentioned above were achieved in tightly controlled laboratory settings with users who often have many hours of training making fine judgments on small changes in thermal stimulation. It may well be that these levels of accuracy do not hold in more realistic, and particularly, mobile settings. Higher rates of change and/or higher stimulus intensities may be required to produce detectable stimuli in these circumstances. For example, Nakashige *et al.* [18] found that not all users could detect changes of +3.9°C or -3.3°C at a rate of ~1°C/sec when the palm of novice users was stimulated in less controlled settings.

Thermal Feedback in HCI

In their summary of thermal perception and the design of thermal feedback for virtual material discrimination in VR, Jones & Berris [12] summarized what they saw as the "desired features" of a thermal display. These were based on both VR research and psychophysical evidence and indicate the range of control a system would need to have to make full use of the thermal sense. They recommend using stimulus temperatures of 22-42°C and employing higher rates of change to maximize detection of stimuli. However, they also suggest a thermal interface should be capable of heating and cooling resolutions of 0.001°C and 0.002°C respectively to mimic the subtle differences in the thermal conductance of different materials. These features are extremely precise, and necessarily so for VR applications, but this level of accuracy may not be necessary or even perceivable for mobile thermal feedback.

Wettach *et al.* [23] designed a Peltier-based thermal feedback apparatus for mobile devices and tested users ability to differentiate three different stimulus temperatures (32°C, 37°C and 42°C). Initially error rates were around 65%, although this number dropped to 25% after long-term training. None of these temperatures would normally be considered 'cool' and so this study suggests that individuals can identify varying degrees of warmth, not simply a change from one temperature to another. This was only an initial study, however, if users can only distinguish these three relatively disparate temperatures at 75% accuracy, it is unlikely they will be able to differentiate stimuli at the resolutions suggested by Jones & Berris [12].

Other work looking at thermal feedback in HCI has been more abstract, often focusing on emotional communication, given the inherent affective quality of thermal sensations, compared to vibrotactile or even auditory output [22]. Gooch [4] found that adding thermal feedback to remote, PC-mediated communication increased feelings of 'social presence'. Nakashige *et al.* [18] accompanied photographs of warm and cold scenes (food such as soup and ice cream, and environmental examples like fire and snow) with either warm or cold feedback. A small informal study indicated that the foods appeared more appealing when accompanied by the corresponding temperature and a small number of users reported an impression of a "loving home" from the warm soup. Iwasaki *et al.* [10] suggested a system that could be used to convey emotional information to another user through augmentation of an existing mobile phone. Lee and Lim [17] discussed existing preconceptions about the meaning or significance of thermal sensations in general and found that users did not treat sensations as binary (i.e. warm and cold) but as a continuum, They also found that sensations were almost meaningless without context but were a very unobtrusive form of feedback. They also found that cold stimuli were generally less preferred than warm.

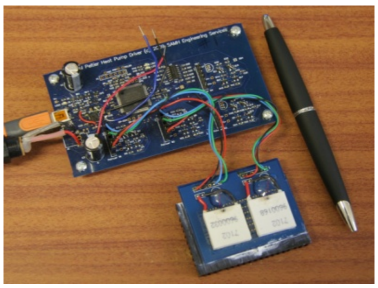

Figure 1: Microcontroller and Peltier stimulators used in the experiments.

Recently, Kushiyama and colleagues [16] have developed thermal display technology utilizing Peltier elements and Narumi *et al.* [19] developed a contextual/ambient thermal display for more mobile scenarios. Neither piece of technology has been tested experimentally but both hold promise for future use. Thermo-pict [16], in particular, stands out as it has the opportunity to provide patterns of thermal stimulation not possible with more limited hardware, including thermal tactile 'pictures' and visible displays.

There are huge differences in feedback design requirements between the highly accurate and articulate thermal interfaces used in VR to represent real materials and the much lower resolution, "warm-or-cold" designs of affective computing. There are many potential uses for thermal feedback in HCI but more work needs to be done on how to design and structure feedback so that it is comfortable, perceivable and suitable for use in a variety of real-world static and mobile settings. As such, we ran two evaluations to identify characteristics of stimulus presentation that would be well suited for use in thermal feedback design for mobile HCI.

EVALUATIONS

We developed a custom microcontroller board that could independently drive four Peltier heat pumps (see Figure 1). Peltier heat pumps allow for a high level of control over temperature output and also allow for both heating and cooling from the same pump. Each Peltier device could be independently controlled over USB, with the temperature set anywhere within the operating temperatures of -20°C to +45°C, accurate to 0.1°C. The Peltiers themselves were bonded to circuit boards with a heat sync bonded to the other side. Safety mechanisms on the board prevented the Peltiers from becoming hotter than 45°C and cut off all power and input if the thermistor detecting the Peltier output became electrically or physically disconnected. For our study we used two Peltiers. Using two instead of one gave a larger stimulus area and meant that we would need lower intensity output to produce detectable sensations [22]. We chose not to use more than two as this would have required too large an area of skin for mobile use.

Static Indoor Evaluation

The initial investigation looked at how well users could detect thermal stimuli while sat at a desk in an indoor usability lab. The Peltier board was controlled over USB from a MacBook Pro running a Pygame interface. User input to the Pygame GUI was received via a mouse.

Stimuli

A neutral starting temperature of 32°C was chosen as this is within the defined 'neutral zone' of thermal sensation [12, 22] and has been used in other studies [5, 14]. The skin was adapted to this temperature before each trial session and was returned to it between each stimulus presentation. Two different rates of stimulus change (ROC) were used: 1°C/sec and 3°C/sec. Previous work has shown that 1°C/sec should be adequate to produce detectable sensations in ideal situations [2, 14, 20]. Other work suggests it may not always be large enough [18] so a higher rate was also included. Not only will this comparison indicate which is the more useful ROC, in terms of producing detectable sensations, but also it would allow for comparison of power requirements, which will need to be balanced with perception performance when being considered for mobile use.

	Warm		Cool	
Intensity	**1°C/sec**	**3°C/sec**	**1°C/sec**	**3°C/sec**
1°C	33°C	33°C	31°C	31°C
3°C	35°C	35°C	29°C	29°C
6°C	38°C	38°C	26°C	26°C

Table 1: Stimuli by intensity, direction and ROC.

Three different stimulus intensities were used: 1°C, 3°C and 6°C. From previous findings, 1°C changes were detectable at rates of change equal to and below those used here, but again in ideal laboratory conditions [2, 14, 20], whereas 3°C was perceivable in a less controlled desktop situation [18]. Wettach *et al.* [23] and Lee and Lim [17] suggest that users were able to differentiate varying degrees of intensity and so a stronger intensity of 6°C was introduced to investigate user responses to varying intensities. As thermal perception is a bipolar sense, both warming and cooling stimuli were used, employing each intensity change in both directions from 32°C neutral. Therefore a single stimulus consisted of warming or cooling at a set intensity (1°C, 3°C or 6°C) at one ROC (1°C/sec or 3°C/sec), for example, warming 3°C at 1°C/sec (see Table1). The temperatures were well away from the cold and heat pain thresholds. Each stimulus in this set was delivered twice, giving a total of 24 stimuli presented at each location (3 intensities x 2 directions x 2 rates x 2 presentations, see Table 1 for details).

As our interest is in using thermal feedback for mobile interaction we selected four body locations that are used to either hold, transport or interact with a mobile device, or locations that could potentially be used for mobile device interaction. The fingers and palm of the hand are the most logical choices, as mobile phones are held against the palm and gripped with the fingers. The thenar eminence (the bulbous area of skin adjoining the thumb) was chosen specifically over the central palm itself due to its apparent increased sensitivity to thermal stimuli compared to the palm [11]. The dorsal surface (hairy skin) of the forearm was chosen partly as it has differing thermal sensitivity to both the thenar eminence and the fingertips [6] but also as it is conceivable that a watch or wrist band could be worn which contains thermal elements. Finally, the upper arm is already used for MP3 players. Such a device could include thermal feedback through skin contact in the same way as one worn around the wrist.

Participants

Fourteen participants (9 male, 5 female) aged 21 to 57 (mean=29.2 years) took part in the evaluation, all studying or working at the University. All were right-handed and paid £6 for participation, which lasted just over an hour.

Variables

For this study we were interested in identifying what stimuli produce detectable sensations from a neutral base temperature, i.e. which stimuli were noticeable and so would be best suited for use in thermal feedback design. The independent variables were: *Rate of change, Stimulus intensity, Direction of change (warm or cool)* and *Body location*. The dependent variables were: *Stimulus detection* (if one was perceived), *Detection time* (how long after the initiation of a stimulus that it was detected), *JND size* (distance in °C from neutral when the stimulus was felt), *subjective intensity of stimulus* and *subjective comfort of stimulus*. We recorded user subjective reports of the intensity of the stimulation (a 7-point Likert scale from “Very Cold” up to “Very Hot”) and the comfort level of the stimulus (a 7-point Likert scale from “Very Uncomfortable” up to “Very Comfortable”) similar to others used before [1, 3].

Environmental temperature influences skin temperature [9, 12] and so is an important factor to consider as particularly high (>25°C) or low (<=15°C) environmental temperatures can cause the skin temperature to shift from the neutral zone. We did not have the facilities to run the experiment in a climate-controlled room and therefore recorded room temperature and humidity during each trial and compared results against these data. It should be noted, however, that the skin under the Peltier was always adapted to the neutral 32°C between trials (and measured to ensure it was so).

Procedure

The task was split up into 4 conditions based on the location of stimulation, with all participants taking part in all conditions in a counterbalanced order and the ordering of stimuli was random. The participant was sat at a desk upon which there was a computer monitor and mouse. For the fingertip and thenar eminence conditions, the Peltier stimulator lay on the desk in front of the seated participant, facing up so that the users could lay their finger/hand on the stimulator, supported by a padded rest (see Figure 2 left). For the forearm and upper arm conditions, the stimulator was held against the arm with an elastic fabric strip secured with Velcro pads. The stimulator was held between this strip and the skin (see Figure 2 right).

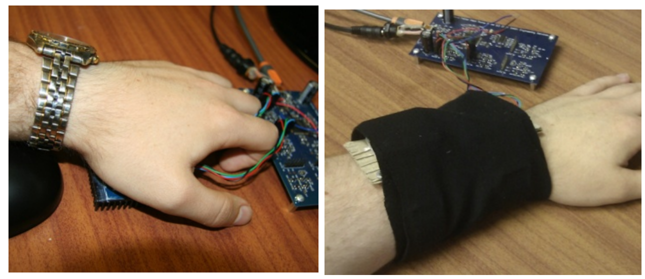

Figure 2: Stimulator sites for thenar and forearm conditions.

The stimulator was contacted with the skin of the non-dominant hand/arm (by resting on top or via the elastic strip) at the beginning of each condition and remained in contact for the duration of that condition. Green [6] found that participants reported higher intensity perceptions when they were in contact with a stimulator between successive stimuli, compared to removing their hand from the stimulator in between trials. At the start of each condition the stimulators were set to the neutral starting temperature of 32°C for one minute so as to adapt the skin to this temperature. After the adaptation period all 24 stimuli were presented in random order. A stimulus presentation comprised of 10 seconds of stimulus followed by a return to the neutral temperature and 20 seconds of adaptation. There were no visual or auditory cues as to when stimuli were presented. Participants were instructed to click the mouse button as soon as they felt a change in thermal stimulation, in any direction and at any intensity. Once this occurred, the temperature of the Peltiers was taken as the temperature that was detected, and the time elapsed since the initiation of the stimulus was taken as the time-to-detection. At this point, 2 Likert scales appeared on screen asking the participants to rate the stimulus felt in terms of *intensity* (from “very cold” to “very hot”) and *comfort* (from “very uncomfortable” to “very comfortable”). They then clicked on a submit button and another stimulus was presented after the 20 seconds of adaptation. If the participant clicked the mouse button before the full 10 seconds of stimulation had passed, the Peltiers were immediately returned to neutral and the rating scales were presented.

Results

Environmental Temperature

During the experiment, room temperature ranged from 20.5°C to 25.3°C with an average of 23.6°C. Humidity ranged from 42% to 67% with an average of 48.5%. Perceptual research has shown that, at room temperatures of 15°C to 25°C, skin temperature sits at neutral temperatures of 30°C to 35°C. So the neutral starting temperature used here would have been close to overall skin temperature, producing no sensation of warm or cool.

Number of Detections

A multi-factorial repeated-measures ANOVA showed a significant effect of body location on the number of thermal detections produced ($F_{3,39}$=3.073, p<.05). *Post hoc* pair-wise Bonferroni-corrected *t*-tests showed the thenar eminence ($\bar{x}$ =87.5%) produced significantly more detections than the fingertips ($\bar{x}$ =75.5%) but non-significantly more than the forearm ($\bar{x}$ =79%) and upper arm ($\bar{x}$ =78.5%).

There was a significant effect of stimulus intensity on the number of detected stimuli ($F_{2,26}$ = 84.623, p<.001). Bonferroni-corrected *t*-test comparisons showed a significant difference between the numbers detected from all stimulus intensities. The number increased as the intensity increased with means of 53%, 90.5% and 97% for 1°C, 3°C and 6°C intensities (see Figure 6 'S' columns). There was no significant effect of ROC or direction of change on the number of stimuli detected. Both warm ($\bar{x}$ =79%) and cool ($\bar{x}$ =81.5%) stimuli produced similar numbers of detections.

Time-to-Detection & JND Size

The time-to-detection and size of JND are directly related and so are considered together. Friedman's analysis of variance by ranks showed location had a significant effect on both time-to-detection (χ^2 (3)=24.71, p<.001) and JND size (χ^2 (3)=41.65, p<.001). Wilcoxon *T* comparisons showed that the finger produced significantly longer times (median=3.22s) than the thenar (median=2.54s; *T*=3672, p=.001), forearm (median =2.68s; *T*=3205, p<.001) and upper arm (median=2.50s; *T*=2988, p<.001). Similarly, Wilcoxon pair-wise comparisons showed that all locations had significantly different JND sizes from each other (p<.05) with the finger (2.9°C) having the largest, compared to thenar (1.9°C), forearm (2.2°C) and upper arm (2.25°C). Significant effects of direction of change showed that warming stimuli were detected significantly slower (median=2.91s) than cooling stimuli (median=2.46s) (*T*=15,643, p <0.001) and, consequently, warm JND size was significantly larger (median=2.80°C) than cold JND size (median=1.85°C) (*T*=8327.50, p<.001).

Stimulus intensity also had a significant effect on time-to-detection (χ^2 (2)=63.01, p<.001; see Figure 3). Wilcoxon *T* comparisons showed significant differences in the time-to-detection between all stimulus intensities: 1°C vs. 3°C (*T*=2906, p<.001), 1°C vs. 6°C (*T*=1638, p<.001) and 3°C vs. 6°C (*T*=9419, p<.01). The amount of time decreased as the intensity increased with median values of 3.67s, 2.59s and 2.30s for 1°C, 3°C and 6°C intensities. Friedman's analysis also showed a significant effect of stimulus intensity on JND size (χ^2 (2)=261.27, p<.001). Wilcoxon comparisons showed that all intensities were significantly different from each other (p<.001) with median JNDs of 1°C, 2.7°C and 3.75°C for intensities of 1°C, 3°C and 6°C respectively.

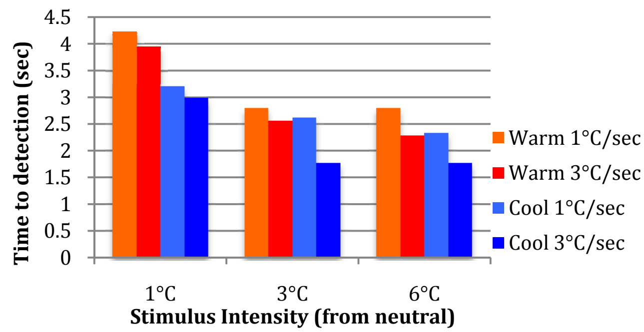

Figure 3: Median detection time at each intensity and rate of stimulus change.

Finally, rate of change significantly affected the time-to-detection (*T*=12,461, p < .001) and JND size (*T*=8213.50, p<.001). The higher ROC (3°C/sec) produced a significantly lower time (median=2.43s) than the lower rate of change (1°C/sec; median=3.04s) while the lower ROC had significantly lower JND size (median=1.90°C) than the higher ROC (median=3.00°C).

Subjective Stimulus Intensity

Intensity ratings ranged from 0 to 3, where 0 denoted "Neutral" and 3 denoted "Very Intense". A Wilcoxon *T* test showed a significant effect of ROC on subjective stimulus intensity (*T*=3935, p<.001). The higher ROC produced significantly higher ratings of intensity (median=1.5) than the low rate of change (median=1.00). There was also a significant effect of direction of change (*T*=7263.50, p<.001) as warm stimuli were rated as significantly more intense (median=1.5) than cold stimuli (median=1.0).

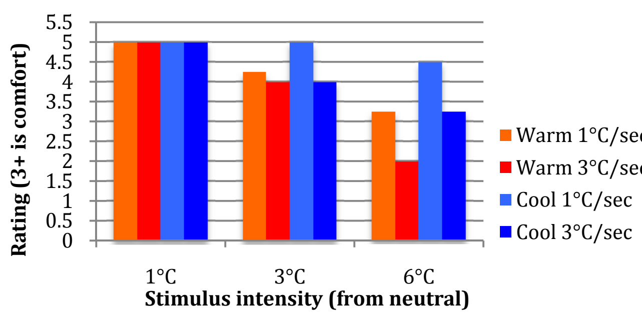

Figure 4: Median subjective comfort ratings at each intensity and rate of change. Rating of 3+ indicates comfort.

A Friedman's ANOVA showed a significant effect of stimulus intensity on perceived stimulus intensity (χ^2 (2)=160.742, p<.001). Wilcoxon *T* comparisons showed that all three stimulus intensities were significantly different from each other (p<.001) with median ratings of 1, 1.5 and 2 for intensities of 1°C, 3°C and 6°C respectively.

Subjective Stimulus Comfort

ROC had a significant effect on subjective comfort (*T*=8775, p<.001) with the higher rate of change (3°C/sec) producing significantly lower ratings of comfort (median=4.0) than the lower ROC (1°C/sec, median=4.50). A Wilcoxon *T* also showed a significant effect of direction of change on reports of comfort (*T*=8922.50, p<.001). Warm stimuli had a significantly lower average comfort rating (median=4.0) than cool (median=4.5), see Figure 4. Friedman's analysis of variance by ranks showed that stimulus intensity also had a significant effect on subjective comfort (χ^2 (2)=48.46, p<.001). Wilcoxon *T* comparisons showed significant differences in subjective comfort between all stimulus intensities: 1°C vs. 3°C (*T*=2107.50, p<.001), 1°C vs. 6°C (*T*=1500.50, p<.001) and 3°C vs. 6°C (*T*=4343.50, p<.001). Ratings of comfort decreased as the intensity increased, with median values of 5.0, 4.0 and 3.0 for 1°C, 3°C and 6°C respectively.

The results of this study show that the thenar is the most sensitive area with the arm locations slightly less so and the fingers the least sensitive. Increasing the rate of change or stimulus intensity makes the stimulus more salient, as it is more quickly detectable and more likely to be detected, but this comes at the expense of subjective comfort. In order to validate these results and also to understand how being mobile influenced stimulus detection, we carried out a follow-up evaluation using similar apparatus and stimuli but with users walking around an indoor environment.

Mobile Indoor Evaluation

For this study, the Peltier microcontroller was connected to an Ultra Mobile Personal Computer (UMPC) running Windows XP, which was carried by each participant in a backpack (see Figure 5). All commands were sent to the UMPC via Bluetooth from a Nexus One phone held in the user's dominant hand, which displayed the experimental interface; these commands were then sent to the Peltier microcontroller over USB via this UMPC. Fourteen participants (10 Male, 4 Female) aged 23 and 41 (mean=30.2 years) took part in the evaluation, all were studying or working at the University. All were right-handed and were paid £10 for participation, which lasted just over an hour.

Procedure

The same stimuli and variables were used as in the static indoor evaluation. However, in this experiment the *bodily location* variable only had 3 conditions. The dorsal surface of the forearm and the dorsal surface of the upper arm were also used in this experiment, as they are common locations where people place devices e.g. a watch on the wrist or an MP3 player on the upper arm. In place of the thenar and fingertips locations, the Peltier devices were placed in the palm of the hand of each participant to simulate holding a mobile phone (and the fingertip condition had poorer performance in the static experiment). The task was split into 3 conditions based on location of stimulation, with all participants taking part in all conditions in a counterbalanced order and the ordering of stimuli was random. Each participant walked around a triangular route in an indoor office environment. For all conditions, the stimulator was held against the arm with an elastic fabric strip secured with Velcro pads. The stimulator was held between this strip and the skin (see Figure 5). The stimulator was contacted with the skin of the non-dominant hand/arm at the beginning of each condition and remained in contact for the duration of that condition. The order and duration of the stimuli were presented in the same manner as the indoor static experiment. Participants were instructed to press a button on the phone screen when they felt a change in thermal stimulation, in any direction and at any intensity. Once this occurred, the temperature of the Peltiers was taken as the temperature that produced the sensation, and the time elapsed since the initiation of the stimulus was taken as the time-to-detection. At this point the same 2 Likert scales as were used in the static experiment appeared on screen and the procedure continued as in the static experiment.

Figure 5: Stimulator locations for forearm and upper arm.

Results

Unfortunately due to issues to do with the external thermistors on the Peltier devices not all conditions were completed for all users. Moisture from skin caused the Peltier devices to stop working when it came into contact with thermistor; this issue was remedied by placing tape over the external thermistors, this still allowed the thermistors to function as normal. As a result 662 out of 1008 stimulations were submitted to the participants. This was taken into account when calculating all of the statistics related to the mobile study. Despite this, statistical tests were still run to compare the results from the static and mobile conditions to empirically test how mobility affected thermal perception.

Environmental Temperature

Room temperature ranged from 20.2°C to 25.9°C, with an average of 21.8°C, and humidity ranged from 39% to 69%, with an average of 48.9%. As with the static experiment the neutral starting temperature used here was close to overall skin temperature, producing no sensation of warm or cool.

Number of Detections

Friedman's analysis of variance by ranks was used to analyse the effect of intensity of change and body location on the number of stimuli detected. Intensity of change was found to have a significant effect on the number of thresholds produced (χ^2 (2)=134.105, p <0.001). *Post hoc* Wilcoxon *T* comparisons showed significant differences in the number of detection between all stimulus intensities: 1°C vs. 3°C (T = 6837.0 p<0.001), 1°C vs. 6°C (T=9100.0 p<0.001) and 3°C vs. 6°C (T=2251.5 p<0.001). Wilcoxon pair-wise comparisons were also used to determine the effect of ROC and direction of change. There was no significant effect of ROC, bodily location or direction of stimulation. The number of stimuli detected increased as the intensity increased with means of 28.44%, 69.68% and 85.78% for 1°C, 3°C and 6°C intensities respectively (see Figure 6). As can be seen, the users are not able to detect as many stimuli as in the static condition but the patterns of detection were approximately the same. When compared with the static results Mann-Whitney *U* comparisons showed that mobility significantly affected the number of stimulus detections (*U*=313515.5, *Z*=12.049, p<.001), with more detections in the static conditions compared with mobile.

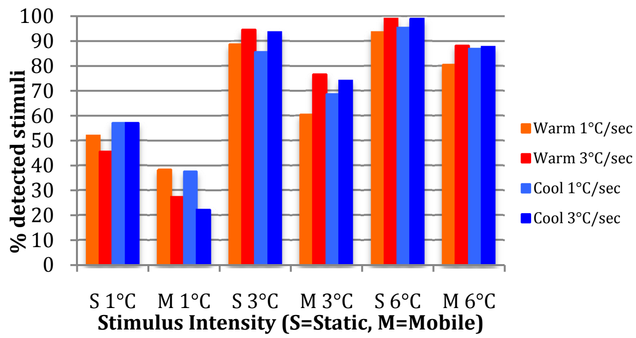

Figure 6: Percentage of stimuli detected at each intensity and rate of stimulus change for both static and mobile studies.

Time-to-Detection and JND Size

Friedman's analysis of variance by ranks was used to analyse the effect of location and stimulus intensity on the average time-to-detection. Location did not have any significant effect, however stimulus intensity did have a significant effect on time-to-detection (χ^2 (2)=36.102, p<.001). Wilcoxon *T* comparisons showed significant differences in the time-to-detection between the smallest intensity and the two larger stimulus intensities: 1°C vs. 3°C (*T*=1687.5, p<.001), 1°C vs. 6°C (*T*=1844.5, p<.001), but not between 3°C and 6°C. The amount of time decreased as the intensity increased with median values of 3.74s, 3.07s and 2.95s for 1°C, 3°C and 6°C intensities respectively. Wilcoxon *T* tests were used to analyse the effect of rate of change and direction of change on time-to-detection. Rate of change significantly affected the time-to-detection (*T*=12,472, p < .001): the higher ROC (3°C/sec) produced a significantly lower time (median=2.96s) than the lower ROC (1°C/sec; median=3.35s). A significant effect of direction of change (*T*=6105.5, p < .001) showed that warming stimuli were detected significantly slower (median=3.28s) than cooling stimuli (median=2.94s).

As stated earlier, time-to-detection is linked directly to the size of JND. A Friedman's ANOVA showed a significant effect of intensity on the size of JND (χ^2 23)=79.993, p<.001). *Post hoc* Wilcoxon pair-wise comparisons showed that all intensities differed significantly from each other (p<0.001) Median JND sizes were 1°C (1°C intensity), 2.3°C (3°C intensity), and 4°C (6°C intensity). Friedman's analysis of variance by ranks showed no significant effect of location on the size of JND. Wilcoxon *T* comparisons showed a significant effect of rate of stimulus change on JND size (*T*=14047.5, p<0.001): 1°C/sec had significantly lower JND size (median=1.90°C) than 3°C/sec (median=3.00°C). A further Wilcoxon *T* also showed a significant effect of direction of change on JND size (*T*=6820, p<0.001). Warm stimuli produced significantly larger JNDs (median=3.1°C) than cold stimuli (median=2.1°C).

When compared with the static results a *U* comparison also showed that time-to-detection was significantly affected by mobility (*U*=33452, *Z*=7.330, p<.001), with mobile detections taking significantly longer than static detections. As threshold size and time-to-detection are interlinked, the size of threshold was also significantly affected by mobility (*U*=69784.0, *Z*=11.440, p<.001), again with a greater threshold size for mobile in comparison with static.

Subjective Stimulus Intensity

As with the static evaluation intensity ratings ranged from 0 to 3, where 0 denoted "Neutral" and 3 denoted "Very Intense". A Friedman's ANOVA showed a significant effect of stimulus intensity on perceived stimulus intensity (χ^2 (2)=160.742, p<.001). Wilcoxon *T* comparisons showed that all three stimulus intensities were significantly different from each other (p<.001) with median/mean ratings of 1/1.09, 1/1.47 and 2/1.95 for intensities of 1°C, 3°C and 6°C respectively. A Friedman's ANOVA showed no significant effect for body location. Similarly, Wilcoxon T tests revealed no significant effect for either stimulation direction or rate of change of intensity.

Subjective Stimulus Comfort

A Wilcoxon T test revealed that ROC had a significant effect on subjective comfort (*T*=7364.5, p<.005) with the higher ROC (3°C/sec) producing significantly lower ratings of comfort (median=4 mean = 3.63) than the lower ROC (1°C/sec, median=4 mean = 3.99). A Wilcoxon *T* test also showed a significant effect of direction of change on reports of comfort (*T*=6408.5, p<.001). Warm stimuli had a significantly lower average comfort rating (median=4 mean=3.52) than cool stimuli (median=4 mean = 3.89). Friedman's analysis of variance by ranks showed that stimulus intensity also had a significant effect on subjective comfort (χ^2 (2)=412.703, p<.005). Wilcoxon *T* comparisons showed significant differences in subjective comfort between 1°C and 3°C (T=1405, p<0.05) stimulus intensities and 3°C vs. 6°C (T=5712, p < 0.001) stimulus intensities. Ratings of comfort decreased as the intensity increased, with median values of 4 (mean =4.1), 4 (mean =4.06) and 3 (mean =3.4) for 1°C, 3°C and 6°C respectively. A Friedman's analysis of body location also revealed a significant effect on subjective comfort (χ^2 (2)=9.968, p<.01). Pair wise Wilcoxon T tests revealed significant differences between the back of the forearm and upper arm conditions (T=39911.5, p<0.005) and between the forearm and palm of the hand conditions (T=3466.5, p<0.01), there were no differences between the upper arm and palm of the hand condition. The median values of the subjective comfort were 4 (mean = 3.63), 4 (mean =3.95) and 4 (mean = 3.71) for the arm, forearm and hand respectively.

DISCUSSION

In order to put these results into context for thermal feedback design for HCI, we will discuss each measure's significance in terms of suitability for use in feedback, either in terms of optimal sensitivity (for body location) or perceivability (for stimulus characteristics ROC, direction of change and intensity). Unless otherwise made clear, all relationships/effects refer to both mobile and static results.

Comparison of Static and Mobile

The results from the mobile study were very similar to those in the static, and as such they go some way to validating both the results themselves as well as the feedback recommendations that arise from them (see this section and Design Recommendations). In general, stimuli were more difficult to detect when mobile, with lower detection rates and slower detection speeds. However, despite the lower absolute values, the static and mobile values changed in very similar patterns. Indeed it was found that mobility significantly affected stimulus detection rates, with far fewer being detected while the user was mobile. We also found that time-to-detection was significantly slower when mobile, compared to static detections. As threshold and time are interlinked the size of threshold was significantly larger for mobile in comparison with static, meaning a larger amount of change occurred before detection.

Another prominent difference between the two studies is the minimal effect body location had on all experimental measures during the mobile study. During the static experiment, body location significantly affected the number of stimuli detected, the time-to-detection and the size of JND, but *not* subjective comfort or intensity. One reason for this difference may be the omission of the finger condition in the mobile study as, comparing the results for the three non-finger conditions in the static study, measures of both number of detections and time-to-detection were very similar. The JND sizes for all locations were significantly different; however the fingertip still performed the worst of all areas. Therefore, it may have been the fingertip performance that skewed the static results. The other important difference between the mobile and static environments was the finding that neither direction of change or rate of change significantly affected subjective intensity while mobile, whereas they both did while stationary. This is discussed in more detail below. Aside from these findings, all other variable relationships/effects were the same across both mobile and static studies and are discussed below.

Location Sensitivity

The measures that relate most to judgments of 'sensitivity' here are the *number of detections*, the *time-to-detection* and the *size of JND*. A higher number of detected stimuli, a faster detection or a smaller JND size are all indications of higher sensitivity. Considering these measures only, the thenar eminence was shown to be the most sensitive location. In all measures in the static study it performed outright best or equal best with high detection rates, low detection speeds and small JND sizes. For the mobile condition the palm was most sensitive, in terms of number of stimuli detected. Both arm conditions performed very similarly to each other, suffering slightly lower detection rates and larger JNDs than the thenar and palm but had roughly equal detection speeds. The finger performed worst out of all locations in all three of these measures. The epidermis of glabrous skin, particularly that on the pads of the fingers, can be up to 5x thicker than that of hairy skin (for example on top of the arm) which increases delay in the thermal stimulus reaching thermoreceptors [24]. Overall detection rates were somewhat lower for the mobile study than when static. Although the hand was more sensitive and both arm locations were roughly equal in both environments, the detection rates were some 10-20% lower when mobile, at 60-65%. Using more intense stimuli when mobile increased detection rates (see Figure 6), up to 86%. From these results we recommend the thenar eminence as the optimal location for thermal feedback, however, in mobile situations where contact with precisely the thenar is not practical, the palm is recommended. Non-glabrous arm locations are also suitable for stimulus presentation. As the volar/glabrous skin on the forearm is suggested to have comparable sensitivity to the thenar [7, 11] this location will also be highly suitable. The fingers are the least suitable location due to low accuracy and slow response time. When the user is mobile stimuli will be more difficult to detect and so more intense feedback is necessary to increase likelihood of detection.

Stimulus Perception

The three measures used for determining location sensitivity are also used to determine how perceivable a stimulus is. However, thermal feedback is said to have an inherent hedonic element [22] and perception often varies from veridical (true) stimulation [21, 22] so it was useful, if not necessary, to also use measures of subjective intensity and comfort in our considerations of a stimulus' suitability for use in thermal feedback. Recall that intensity values here run from 0 (neutral) to 3 (very strong) and comfort ratings ranged from 0 (very uncomfortable) up to 6 (very comfortable) with a neutral value at 3. We therefore considered any rating of 3 or above as indicating acceptable comfort levels for feedback design.

Rate of Stimulus Change

Previous research has suggested that stimulus detection is heavily influenced by the rate of stimulus change, with higher rates of change producing more noticeable stimuli [2, 5, 8, 14]. This position was only partly supported by the results here, as, although the high ROC (3°C/sec) produced a significantly faster time-to-detection, both rates produced equal numbers of detections, which suggests higher ROC did *not* produce more noticeable stimuli. Given the lower detection times, it seems that ROCs over 1°C/sec do not affect *if* a stimulus is detected as much as *when* it is detected. The faster time suggests that faster-changing stimuli are more salient, or more *immediately* salient, and so thermal interactions that are time-critical would be recommended to use faster ROC to bring attention to the event more quickly. We also found that the high ROC produced significantly larger JND sizes and was significantly less comfortable than the low ROC. Some work has shown that JND size decreases as ROC increases up to about 3°C/sec at similar body locations [2] so this result is slightly puzzling. Further, the high ROC felt subjectively more intense than the low ROC in the static setting, but not while mobile. Temporal summation would cause faster changes in a small amount of time to be perceived as being more intense [22] which would explain both the higher intensity ratings while stationary as well as the lower comfort ratings (due to higher intensity). Why the high ROC was not more intense and yet still less comfortable in the mobile condition is unclear. In terms of feedback design, it appears that increasing the rate of change brings no benefit to stimulus detectability, and reduces stimulus comfort; however it can be used to reduce detection time and to potentially increase subjective intensity of the stimulus.

Direction of Change

In almost all measures, cold stimuli were more perceivable than warm stimuli. Although they produced a roughly equal number of detections, cold stimuli were faster to detect, produced smaller JNDs and were more comfortable. Cold perception has been found to be faster than warm [12, 20], producing smaller JNDs [14], and this was also found here. These three factors (time, JND and comfort) make cold stimuli particularly appealing for feedback design as they require less power to produce a detectable amount of change and are detectable sooner, compared to warm stimuli. Warm stimuli were reported as more subjectively intense than cold during the static condition but not while mobile. The human neutral thermal zone is around 28-40°C [12, 22], however the relative proximity of heat-pain and cold-pain to this area is not equal (45°C and 15°C respectively [20, 22]). It therefore takes more cooling to feel cold-pain than it does warming to feel heat-pain. In this study, the highest temperature used (38°C) was closer to heat-pain than the coldest temperature (26°C) was to cold pain [22]. Therefore, as the warm stimuli here were closer to painful levels, it may well be that these felt more intense than cold stimuli. What this means for feedback design is that care must be taken not to use too hot stimuli, as these may be too intense. Not surprisingly, these more intense warm stimuli also felt less comfortable. Further, more extreme cold will be necessary to produce equally intense stimulations as intense warmth, requiring disproportionately more power.

Stimulus Intensity

Here "intensity" refers to the maximum amount of change from neutral, set at 1°C, 3°C and 6°C. Although most research on thermal perception focuses on the rate of change rather than the 'end-point' intensity, we chose to control and manipulate stimulus intensity here, as we were interested in designing structured feedback for HCI, and so wished to indentify set stimulus characteristics that produce detectable stimuli. Previous psychophysical work has suggested that 1°C stimuli are highly detectable at lower rates of change than those used here [5, 14], although we hypothesised that such a low intensity may not be suitable for less controlled experimental situations with less highly trained users, which seems to be the case here. The results showed that increasing intensity significantly increased the number of stimuli detected, and as can be seen from Figure 6, detection rates for 1°C stimuli were very low, compared to those of 3°C and 6°C. Therefore, it seems that 1°C changes are not suitable for thermal feedback due to the low detection rate. Increasing stimulus intensity also significantly increased JND size and significantly reduced time-to-detection. The significant reduction in time-to-detection was mainly limited to between 1°C and 3°C intensities, which would suggest that the extra power required to push further change to 6°C would be wasted, purely in terms of speed of detection. Perhaps unsurprisingly, each intensity felt significantly more intense than those beneath it. This is a useful finding as it suggests that, in agreement with Wettach *et al.* [23], participants can differentiate at least some discrete levels of warmth and cold, not simply detect dichotic 'warming' or 'cooling'. This is an important result for feedback design as it suggests intensity can be used as a design parameter or event mapping. It also means that, although the extra power required for more intense stimuli does not bring about faster detections, it can produce stimuli of perceptually distinct intensity, with increased likelihood of detection.

DESIGN RECOMMENDATIONS AND CONCLUSIONS

From the results we have highlighted several interesting and important factors that should be considered when designing thermal feedback both in general and for mobile devices.

The thenar eminence is the optimal location for feedback, but non-glabrous arm locations are also suitable. In measures of 'number of stimuli detected', 'time-to-detection' and 'size of JND' the thenar eminence either performed outright best or equal best showing it to be the most sensitive area. Similarly, the palm was superior when mobile. Although the forearm and upper arm suffered lower overall detection rates than the thenar/palm, they both performed well on measures of 'time-to-detection' and 'JND size' indicating suitability. The fingers performed poorly on all measures, showing them to be slow and inaccurate in thermal perception.

1°C/sec and 3°C/sec changes are suitable, even necessary, but power-requirements must be considered. Both rates of change produced approximately equal numbers of detectable stimuli, with the best detection rates when using 3°C and 6°C stimuli. Each has its own advantages, however. 1°C/sec changes are slower and require a larger change to detect but feel less intense and so feel more comfortable. Therefore, low rates of change may be best suited to ambient displays. 3°C/sec changes, however, are much faster to detect but sacrifice a degree of comfort without any benefit in likelihood of detection. Faster changes may be necessary when mobile, as they will increase likelihood of detection.

Warm and cold stimuli are both suitable for use. Although both warm and cold stimuli are equally detectable, cold stimuli are faster to detect, require less change to detect and are more comfortable as they feel less intense. Warm stimuli should be used carefully or more subtly as they are generally less comfortable and feel more intense. A potentially problematic effect of using Peltier-based apparatus for producing cold feedback is that of heat sinking. As the skin side cools, the other warms, which could then potentially increase the temperature within the housing/body of the device. Drawing heat away from this side of the Peltier would be necessary for safe and effective use.

Intensity or 'end-point' can be used as a parameter for feedback design as different intensities are perceptually different. This would allow for many levels of warm and cold to be used for event semantics, e.g. priority for messages, proximity for GPS or download status. 1°C intensities are best avoided despite their low power costs and high comfort level, due to high levels of missed stimuli and slow detection speed. Finally, high intensities are best used at lower rates of change, as this will minimise discomfort.

Further work will study thermal perception of the same stimuli when a user is in a more realistic outdoors environment, as it is necessary to understand how environmental factors influence perception. In conclusion, this paper presented two studies, which investigated how well users were able to detect warm and cold stimuli presented to four locations on the hand and arm with a view to identifying features of stimulus presentation that would be well suited for use in thermal feedback design for HCI. Simple guidelines for design of feedback based on our results are outlined, with attention paid to perceptual and hedonic factors as well as more practical concerns such as power requirements.

ACKNOWLEDGEMENTS
This research has been funded by the Industrial Members of MobileVCE (www.mobilevce.com), with additional financial support from EPSRC grant EP/G063427/1.

REFERENCES

1. Cherycroze, S. Painful Sensation Induced by a Thermal Cutaneous Stimulus. *Pain*, 1983. 17(2): p. 109-137.
2. Claus, D., M. Hilz, I. Hummer, and B. Neundorfer. Methods of measurement of thermal thresholds. *Acta Neurol Scand*, 1987. 76(1): p. 8.
3. Gagge, A., J. Stolwijk, and J. Hardy. Comfort and Thermal Sensations and Associated Physiological Responses at Various Ambient Temperatures. *Environ Res*, 1967. 1(1): p. 1-20.
4. Gooch, D. An Investigation into Communicating Social Presence With Thermal Devices. *MSc Dissertation*, 2009: p. 1-390.
5. Gray, L., J. Stevens, and L. Marks. Thermal Stimulus Thresholds - Sources of Variability. *Physiol Behav*, 1982. 29(2): p. 355-360.
6. Green, B.G. Temperature perception on the hand during static versus dynamic contact with a surface. *Atten Percept Psycho*, 2009. 71(5): p. 1185-1196.
7. Hagander, L., H. Midani, and M. Kuskowski. Quantitative sensory testing: effect of site and skin temperature on thermal thresholds. *Clinical neurophysiology*, 2000. 111(1): p. 5.
8. Harrison, J. and K. Davis. Cold-evoked pain varies with skin type and cooling rate: a psychophysical study in humans. *Pain*, 1999. 83(2): p. 123-135.
9. Hirosawa, I., H. Dodo, M. Hosokawa, S. Watanabe, K. Nishiyama, and Y. Fukuichi. Physiological Variations of Warm and Cool Sense with Shift of Environmental-Temperature. *Int J Neurosci*, 1984. 24(3-4): p. 281-288.
10. Iwasaki, K., T. Miyaki, and J. Rakimoto. AffectPhone: A Handset Device to Present User's Emotional State with Warmth/Coolness. *BIOSTEC 2010*, 2010: p. 1-6.
11. Johnson, K., I. Darian-Smith, and C. LaMotte. Peripheral neural determinants of temperature discrimination in man: a correlative study of responses to cooling skin. *Journal of Neurophysiology*, 1973. 36(1): p. 24.
12. Jones, L.A. and M. Berris. The Psychophysics of Temperature Perception and Thermal-Interface Design. In *HAPTICS '02*, Orlando, FL, 2002.
13. Kenshalo, D. Somesthetic sensitivity in young and elderly humans. *Journal of Gerontology*, 1986. 41(1): p. 11.
14. Kenshalo, D., C.E. Holmes, and P.B. Wood. Warm and Cool Thresholds as a Function of Temperature Change. *Perception & Psychophysics*, 1968. 3(2A): p. 4.
15. Kenshalo, D. and H. Scott. Temporal course of thermal adaptation. *Science*, 1966. 151(1): p. 2.
16. Kushiyama, K., T. Baba, K. Doi, and S. Sasada. Thermo-Pict neo. *SIGGRAPH 10*, 2010: p. 1-1.
17. Lee, W. and Y.K. Lim. Thermo-Message: Exploring the Potential of Heat as a Modality of peripheral Expression. In *CHI 10: Extended Abstracts*, Atlanta, GA, 2010. ACM Press.
18. Nakashige, M., M. Kobayashi, and Y. Suzuki. "Hiya-Atsu" media: augmenting digital media with temperature. In *Proc. of CHI '09*, Boston, MA, 2009.
19. Narumi, T., T. Akagawa, Y.A. Seong, and M. Hirose. Thermotaxis. *SIGGRAPH 09*, 2009: p. 1-1.
20. Pertovaara, A. and I. Kojo. Influence of the rate of temperature change on thermal thresholds in man. *Exp Neurol*, 1985. 87(1): p. 7.
21. Stevens, J. and L.E. Marks. Spatial Summation of Cold. *Physiology and Behaviour*, 1979. 22(1): p. 7.
22. Stevens, J.C., *Thermal Sensibility*, in *The Psychology of Touch*, M.A. Heller and W. Schiff, Editors. 1991, Lawrence Erlbaum: New Jersey.
23. Wettach, R., C. Behrens, A. Danielsson, and T. Ness. A thermal information display for mobile applications. In *Proc. of MobileHCI '07*, Singapore, 2007. ACM Press.
24. Whitton, J. and J. Everall. The thickness of the epidermis. *British Journal of Dermatology*, 1973. 89: p. 10