"Baby It's Cold Outside": The Influence of Ambient Temperature and Humidity on Thermal Feedback

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

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

ABSTRACT
Thermal feedback is a new area of research in HCI and, as such, there has been very little investigation of the impact of environmental factors on its use for interaction. To address this shortcoming we conducted an experiment to investigate how ambient temperature and humidity could affect the usability of thermal feedback. If environmental conditions affect perception significantly, then it may not be suitable for mobile interactions. Evaluations were conducted outdoors in varying environmental conditions over a period of 5 months. Results showed that the ambient temperature has a significant impact on people's ability to detect stimuli and also their perception of these stimuli. Humidity has a negligible effect for most humidity values. Despite this, previous thermal feedback design recommendations still hold in varying temperatures and humidities showing that thermal feedback is a useful tool for mobile interaction.

Author Keywords
Thermal feedback; mobile interaction; environment; temperature; humidity, ambient; non-visual feedback

ACM Classification Keywords
H.5.2 [Information Interfaces And Presentation]: User Interfaces - Haptic I/O;

General Terms
Design, Human Factors

INTRODUCTION
Physical interactions such as multitouch and 3D spatial input through gestures are becoming more prominent in modern mobile devices. While the use of haptic feedback is common and has received a lot of attention, this has mostly been through the vibrotactile actuators built into phones. Other aspects of touch can be exploited to enhance interaction with mobile devices. In particular, in this work we focus on thermal feedback which has a number of potential benefits. It can act as a non-visual notification channel for situations that are too bumpy or noisy for vibrotactile or audio feedback. It can enhance both visual and non-visual feedback by adding extra depth to the interaction experience, e.g. thermal feedback could be used to add affect that is not provided by other modalities [11, 18]. Thermal output is also entirely private; audio may be heard by others, vibrotactile may be heard and felt by others and visual may be seen by others. Currently, the basis for designing thermal feedback comes from the wealth of research on temperature perception from the psychology literature, where the goals are very different, e.g. the investigation of the characteristics and limits of perception, pain thresholds etc.

No previous work in HCI investigating thermal feedback has taken environmental factors, such as ambient temperature or humidity, into account. This means it is not clear if the same thermal feedback could be used in Northern Europe during winter and in the Middle East during summer. If the wrong feedback is delivered in the wrong environmental conditions then the thermal feedback may become ineffective, e.g. warm feedback in hot and humid conditions may be uncomfortable and difficult to use. Thus, it is important to investigate these contextual factors to design effective and usable thermal feedback. With the goal of examining the impact of ambient temperature and humidity on an individual's ability to detect and use thermal feedback, we present a study into how well users perceive hot and cold stimuli on the hand and wrist. Evaluations were carried out outdoors in varying environmental conditions over a period of 5 months. In addition, we examine if any impact of environmental conditions can be overcome with training. The results will help us understand how to construct thermal feedback for a user interface that is usable in a wider range of usage contexts.

BACKGROUND
Temperature Perception
The skin rests in a narrow 'neutral' zone, ranging from ~28°C up to ~40°C when in all but the most extreme environmental conditions [13]. The size of this zone is relatively constant across individuals, at around 6-8°C, but the relative position of each individual's neutral zone varies. Within the neutral zone there is no discernible thermal sensation [13]. Outside of this range a constant sensation of warmth (above) or cold (below) is perceived [15]. Thermal *thresholds* (the amount of temperature change required before a stimulus becomes perceivable, measured from a set baseline temperature) are inextricably linked to both the skins current temperature and the rate of stimulus change

(i.e. rate of warming or cooling). There is a roughly U-shaped relationship between rate of change (ROC) and thermal threshold, in ideal laboratory settings using highly trained individuals. Thermal thresholds decrease to as small as 0.5°C as ROC increases from 0.01°C/sec to 3°C/sec. They then increase again under faster changes, due to conduction velocity as well as reaction and cognition time [9]. This would suggest that, as ROC increases, stimuli become more salient, however, above a set speed, even if salience increases, the ROC 'over-takes' 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 extreme of the neutral zone, warm and cold thresholds, respectively, decrease and decrease more as the stimulus intensity approaches the heat/cold pain thresholds (~45°C and 11-15°C respectively [12, 20]). Conversely, warm and cold thresholds 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 the temperature itself.

Thermal sensitivity is not uniform; there are marked variations between different body locations and different skin types. Glabrous skin (hairless skin as found on the fingertips, palm and volar/under surface of the arm) is generally less sensitive than non-glabrous/hairy skin with normal and pain thresholds generally larger on glabrous skin [9]. This is due to the thickness of glabrous skin compared to hairy skin [9]. 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 [9]. In general, thermal sensitivity is best on the head and trunk but worse towards the extremities [7].

Influence of Environment and Context

A number of studies outside of computing have investigated the effect of the environment on thermal perception and pain thresholds. Strigo *et al.* [21] investigated the effect of ambient temperature on stimulus intensity, pain intensity and unpleasantness of thermal stimuli applied to the volar forearm. The ambient temperatures they used were 15°C (cool), 25°C (neutral) and 35°C (warm) and their thermal stimuli was in the range 0-50°C. They found that mean skin temperature, heat perception (44-50°C) and cold perception (0-25°C) were affected by ambient temperature. They also found that cold stimuli were rated as being more unpleasant in cold conditions, with warm stimuli being more unpleasant in warm conditions. In related work, Gagge *et al.* [3] investigated the effect of ambient temperature on comfort and thermal sensation for seated unclothed subjects. The ambient temperatures they investigated were in the range 12-48°C. They found neutral temperatures to be in the range of 28-30°C. They also found that discomfort increased more rapidly for cool (<28°C) than warm (>30°C) environments, while thermal sensation increased outside the neutral zone. Discomfort increased for cool stimuli in cold environments and for warm stimuli in hot environments. In general, discomfort was linked to a change in average body temperature from 36.5°C. Hagander *et al.* [7] reported that warm and cool pain thresholds were independent of local skin temperature, when temperatures range between 27-37°C. Hirosawa *et al.* [10] found that warm and cold thresholds at the middle fingertips were affected by room temperature, from a neutral zone of 15-25°C.

With respect to humidity, Givoni *et al.* [4] report the results of 5 comfort studies to investigate the impact of humidity on perceived levels of comfort. The 5 studies cover both indoor and outdoor conditions and were conducted in a number of countries (Thailand, Singapore, Indonesia and Japan). They report the effect of humidity on perceived comfort to be small to negligible. They note that the participants were acclimatised to the environment of the country in which the evaluation was conducted.

Although not investigating ambient temperature and humidity, Halvey *et al.* [8] examined how clothing might affect usability of thermal feedback. They studied how well users perceive hot and cold stimuli on the hand, thigh and waist. Evaluations were carried out with cotton and nylon between the thermal stimulators and the skin. Results showed that the presence of clothing required higher intensity thermal changes for detection but that these changes were more comfortable than direct stimulation on the skin.

Most of the studies mentioned above were conducted in tightly controlled laboratory settings with users who often had many hours of training in making fine judgments on small changes in thermal stimulation. It is important to investigate these phenomenons and their application to HCI in more realistic usage scenarios.

Thermal Feedback in HCI

In their review of thermal perception and the design of thermal feedback, Jones and Berris [13] summarised 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 so as 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 (their work was on discrimination of materials in VR environments). These features are extremely exact, and necessarily so for the VR application, however this level of accuracy may not be necessary or even useful for everyday indoor and outdoor mobile thermal feedback needs.

Wettach *et al.* [22] designed a Peltier-based thermal feedback apparatus for mobile devices and tested users' ability to differentiate three 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. 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 and Berris [13]. Wilson *et al.* [23] presented two studies into how well users could detect hot and cold stimuli presented to the fingertips, the palm, the dorsal (upper) 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 was most sensitive, cold was more perceivable and comfortable than warm and that stronger and faster-changing stimuli were more detectable but less comfortable.

Gooch [5] found that adding thermal feedback to remote, PC-mediated interpersonal 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. Both Iwasaki *et al.* [11] and Fujita and Nishimoto [2] have suggested systems that could be used to convey emotional information to another user through either augmentation of an existing mobile phone [11] or through a simple device worn around the neck. Lee and Lim [17] discussed existing preconceptions about the meaning or significance of thermal sensations in general. They found that users did not treat sensations as binary (i.e. warm and cold) but as a continuum; sensations were almost meaningless without context and that it was a very unobtrusive form of feedback. They also found that cold stimuli were generally less preferred than warm.

Recently, Kushiyama and colleagues [16] developed thermal display technology utilising Peltier elements and Narumi *et al.* [19] developed a contextual/ambient thermal display for more mobile scenarios. Neither piece of technology has been evaluated 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 'pictures'.

There are large differences in feedback design requirements for VR, abstract uses of thermal output or for highly controlled lab settings. For use in real world applications more robust evaluations of thermal output must be conducted. Many of the studies looking at the impact of temperature and humidity were conducted with different goals to ours, e.g. they focus on pain thresholds or are conducted in highly controlled environments. Therefore, we conducted a long term evaluation to investigate the impact of ambient temperature and humidity on thermal feedback usability.

EVALUATION

The main goal of this evaluation is to investigate the impact of environmental factors on thermal feedback. There are a number of specific research questions that we investigated:

1. Wilson *et al.* [23] put forward a number of design recommendations for use of thermal interfaces for mobile interaction, namely: (1) the Thenar eminence is the optimal location for feedback, but non-glabrous arm locations are also suitable, (2) 1°C/sec and 3°C/sec changes are both usable and (3) Warm and cool stimuli are both suitable for use, although users have a preference for cool. Their recommendations were based on indoor conditions. We wanted to examine if these recommendations still held in different environmental conditions?
2. Does ambient temperature have an impact on people's ability to detect and perceive thermal stimuli?
3. Does ambient humidity have an impact on people's ability to detect and perceive thermal stimuli?
4. Can repeated use of thermal interfaces help users to overcome any influence of temperature and humidity?

Equipment

For our evaluation we used a microcontroller connected to a four-channel Peltier heat pump designed and built by SAMH Engineering [23] (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. All four Peltier devices could be independently controlled over USB, with the temperature set anywhere within the range of -20°C to +45°C, accurate to 0.1°C. The Peltiers themselves were bonded to circuit boards and therefore it was necessary to bond heat sinks to an exposed copper strip on the underside of the boards to help dissipate this heat.

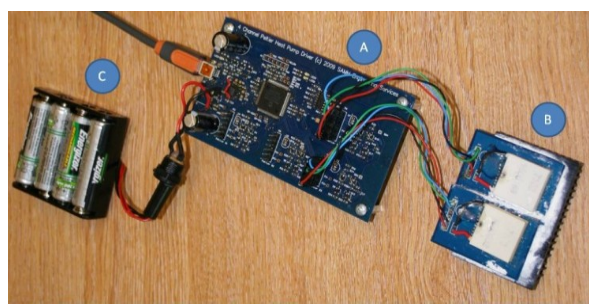


Figure 1: Microcontroller (A), battery pack (C) and Peltier stimulators (B)

Although our device can control four Peltier pumps, only two were used for this study, placed next to each other to stimulate adjoining areas of the skin. The larger the area of stimulation, the greater the thermal sensation, known as spatial summation [20]. Thus using two Peltiers instead of one (i.e. having a larger area of stimulation) meant that we needed lower intensity output to produce threshold sensations, compared to using one. We chose not to employ more than two, even though this could have reduced the intensity requirements further, as this would have required too large an area of skin to fit with the premise of mobile use.

Stimuli

We replicated the design of the study of Wilson *et al.* [23] to get a baseline for indoor performance, so the thermal stimuli used were the same as theirs. Using similar stimuli to previous evaluations allows us to compare our results directly with results of other experiments. A neutral starting temperature of 32°C was chosen as this is within the defined 'neutral zone' of thermal sensation [14, 23]. 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 were used: 1°C/sec and 3°C/sec. Three different stimulus intensities were used: 1°C, 3°C and 6°C. Employing each intensity change in both directions from 32°C neutral gave absolute stimulus temperatures of (in increasing order): 26°C (-6°C), 29°C (-3°C), 31°C (-1°C), 33°C (+1°C), 35°C (+3°C) and 38°C (+6°C). Every temperature was away from the cold and heat pain thresholds (approximately 11-15°C and 45°C respectively). 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).

As our interest is in mobile interaction, we selected 2 body locations that could be used for thermal feedback: the Thenar eminence and the back of the wrist. An additional reason for choosing these locations is that they are often exposed and not covered by clothing. Thus they are exposed to the environment and our evaluations would not be hindered by participants wearing different clothing in different environmental conditions. The Thenar eminence was chosen specifically over the central palm itself due to its increased sensitivity to thermal stimuli [23]. If a person was to hold a mobile device in his/her hand it would be in direct contact with the Thenar. The wrist/dorsal surface (hairy skin) of the forearm was chosen as it has differing thermal sensitivity to the Thenar eminence [6, 23] but also as it is conceivable that a watch or wristband could be worn which forms part of an interaction paradigm for mobile devices. The volar surface (underside) of the forearm was specifically not chosen as research has shown it has equivalent sensitivity to the Thenar eminence [7].

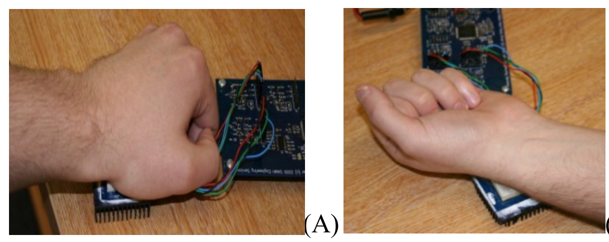
Variables

For this study we were interested in identifying what stimuli produce threshold sensations from a neutral base temperature, i.e. which stimuli were noticeable and so be best suited for use in thermal feedback design. The independent variables were: (1) rate of change (ROC), (2) stimulus intensity, (3) direction of change (warm or cool), (4) body location and (5) user experience with thermal feedback.

The dependent variables were: (1) threshold perception (if one was perceived), (2) threshold time (how long after the initiation of a stimulus that the threshold was perceived), (3) subjective comfort of stimulus and (4) subjective intensity of stimulus. Threshold size (distance in °C from neutral when the stimulus was felt) was also considered as a dependent variable but was removed as it is correlated with threshold time i.e. given a rate of change and time to detection we can calculate threshold size. 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, 21, 23]. For analysis, the intensity scales were mapped to a 0-3 scale, with 0 being neutral and 3 being very hot or very cold for the hot or cold stimuli.

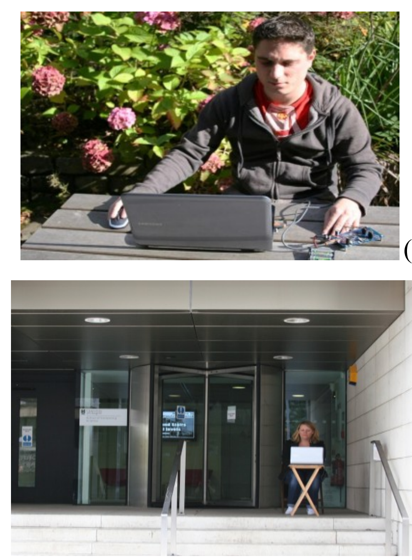
Procedure

The task was split into 2 conditions based on the location of stimulation, with all participants taking part in both conditions in a counterbalanced order. Each participant was seated outdoors at a desk upon which there was a laptop and mouse. The Peltier stimulator lay on the desk in front of the participant, facing up so that he/she could lay a hand or wrist on the stimulator (see Figure 2).



(A) (B)

Figure 2: Stimulator locations for Thenar (A) and forearm (B) conditions



(A)

(B)

Figure 3: Locations for evaluation, (A) garden area and (B) covered entrance way

Two different outdoor locations were used for the experiment to provide a robust and realistic test environment. Both were sheltered to reduce effects of wind. The first was in a small garden area (see Figure 3A), however when it was raining it was not possible to use this location and a covered entrance way was used (see Figure 3B). The garden area was preferred, as the shading in the entrance way meant that we might not get a large range of temperatures.

The stimulator was contacted with the skin of the non-dominant hand/arm (by resting on top) 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 between trials. At the start of each condition the Peltiers were set to the neutral starting temperature of 32°C for two minutes so as to adapt the skin to this temperature. After the two minutes of adaptation 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 30 seconds of adaptation. There were no visual or auditory clues as to when stimuli were presented.

Participants were instructed to click the left mouse button as soon as they felt a change in thermal stimulation, in any direction and at any intensity. Once this occurred, the time elapsed since the initiation of the stimulus was taken as an indication of the threshold time. At this point the two Likert scales appeared on screen asking the individual to rate the stimulus in terms of intensity and comfort. They then clicked on a submit button and another stimulus was presented after the 30 seconds of adaptation had completed. If the participant clicked the button before the full 10 seconds of stimulation had passed, the Peltiers were immediately returned to neutral and the following scale ratings corresponded to the preceding stimulus. If a stimulus was not felt, and so ran its 10 second presentation with no click from the participant, the Peltiers were then returned to neutral and the 30 seconds of adaptation began. In this case, the participant may have felt the transition back to neutral and so a click during this period produced exactly the same scales and data as before, however they corresponded to the transition back to neutral and not the initial stimulus.

Participants

Two groups of users participated: single session users and repeated users. Single session users participated in one session of the study only. Repeated users took part in a session once a month for the five months of the study. Each month six single session users took part, resulting in a total of 30 single session users. Six repeated users started the evaluation; however one had to withdraw, leaving a total of five who completed the whole study. One month was chosen as the time between sessions to provide a greater chance of different environmental conditions (see Figure 4 for example environmental conditions over the period of the study).

The single session group consisted of 18 males and 12 females aged 21 to 39 (mean = 28.77, median = 29), the majority of whom were studying or working at the University. 28 were right-handed and paid £6 for participation, which lasted just under an hour. The repeated session group consisted of 5 males aged 24 to 31 (mean = 26.2, median = 24), all studying or working at the University. All were right-handed and paid £6 per session. The majority of participants were of Northern European origin, although there were individuals from Asia, Africa, North America and South America. All were living in Northern Europe and in that respect can be considered acclimatised.

Environmental Conditions

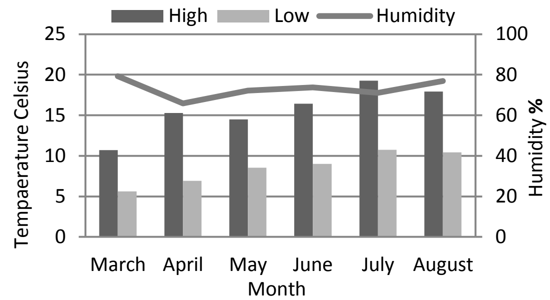
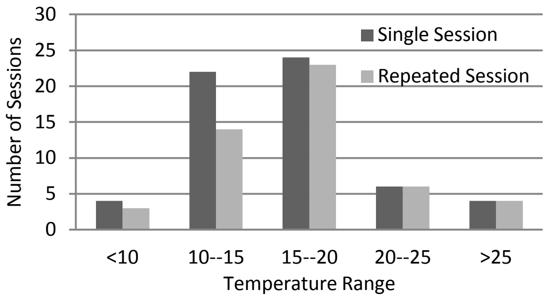


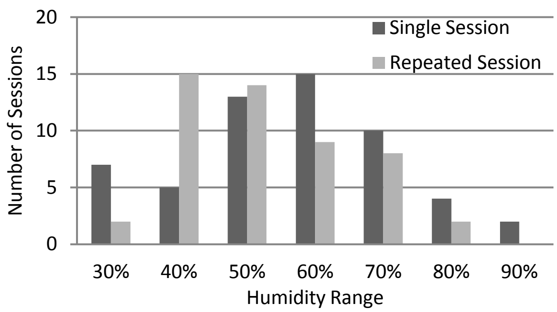
Figure 4: Environmental conditions over 5 months of the study (August is included as some of the July participants finished in the first few days of August)

Participants were tested from March to July (some of the July participants took part in the evaluation in the first few days of August). During each trial, the ambient temperature and humidity were recorded. Temperature and humidity readings were taken 3 times during the evaluation, as over the period of the evaluation these conditions could change, allowing us to calculate the average ambient temperature and humidity. Temperature was measured to an accuracy of 0.1ºC and humidity to an accuracy of 0.1%. Figure 4 shows the average daily high and low temperature and average daily humidity per month for the duration of the experiment. These figures were collected daily from local weather reports for the duration of the evaluation. Initially, we had considered beginning the experiment when temperatures were cooler, however, for ethical, safety and comfort reasons we did not want to have participants seated outside in cold temperatures.

Over all of the evaluation sessions, temperature ranged from a minimum of 8.45ºC to a maximum of 27.75ºC, with humidity from 31.4% of to a maximum of 93.2%. Due to the range of environmental conditions, the data was grouped based on temperature and humidity. Temperature data was grouped in blocks of 5ºC. Previous work by Hirosawa *et al.* [10] found that the neutral zone for thresholds is in the range 15-25ºC. The grouping chosen maintains this neutral zone, while at the same time allows a finer grained analysis of the data. Had we used the 10ºC gap for the 15-25ºC neutral zone, we would only have had three groups; the larger number of groups still allows us to conduct the same analysis. This resulted in 5 temperature based groups <10ºC, 10-15ºC, 15-20ºC, 20-25ºC and >25ºC. Where for example the 10-15ºC range includes temperatures starting at 10ºC up to 15ºC, so there was no overlap between ranges. Humidity data were grouped in blocks of 10%, resulting in 7 blocks: 30%, 40%, 50%, 60%, 70%, 80% and 90%. Givoni *et al.* [4] reported that humidity has a negligible effect on comfort, as such we believed that this grouping would allow us to investigate this phenomenon while still be fine grained enough to show any effects. Figure 5 shows the breakdown of user evaluation sessions by temperature and humidity ranges. While Figure 4 shows the seasonal changes over the course of the experiment, Figure 5 shows the fluctuations for sessions involving experimental participants. As can be seen many of the sessions were conducted in the mid-range temperature and humidities.



(A)



(B)

Figure 5: Experimental sessions broken down by temperature (A) and humidity (B) ranges. Session refers to one body location per evaluation participation.

RESULTS

Single Session

In this section we present an analysis of the factors for the single session participants. The independent variables were analysed as follows: a Friedman's analysis of variance by ranks was used to analyse the effect of ambient temperature, humidity and intensity of change. Wilcoxon pair-wise comparisons were also used to determine the effect of location, ROC and direction of change.

Number of Detections

Table 1 shows the performance for the different ambient temperatures and it can be seen clearly that detection rate peaks in the 15-20°C temperature range. Indeed ambient temperature was found to have a significant effect on the number of thresholds produced (χ^2 (4)=68.325, p<0.001). In addition there are significant differences between a number of the temperature bands (as shown in Table 2). This indicates that effect of ambient temperature might be more fine grained than indicated by previous research [10]. Table 3 shows the figures for the impact of humidity on user performance. Humidity was found to have a significant effect on the numbers of thresholds produced (χ^2 (6)=28.74, p<0.001). It can be seen from viewing the figures in Table 3 that the detection rates are mostly in a similar range with the exception of extreme humidities (i.e. >90% and <30%). *Post hoc* Wilcoxon T comparisons between all humidities showed that the differences between the extreme humidities and more central humidities were significant (figures not shown for reasons of space). While some other differences were found there was no distinct pattern other than differences at the extreme humidity values, in keeping with the findings of Givoni *et al.* [4].

	<10°C	10-15°C	15-20°C	20-25°C	>25°C
Detection Rate (Accuracy %)	82.29 σ=38.37	74.7 σ=43.48	84.25 σ=36.44	79.84 σ=40.27	35.56 σ=48.13
Detection Time (Sec.)	3.46 (3.36) σ=1.37	3.23 (2.82) σ=1.70	3.03 (2.68) σ=1.52	4.14 (3.65) σ=2.25	3.7 (3.93) σ=1.92
Comfort (Scale 0-6)	3.11 (3) σ=1.05	2.09 (2) σ=1.53	1.82 (1) σ=1.61	2.71 (2) σ=1.60	2.38 (2) σ=1.45
Intensity (Scale 0-3)	1.24 (1) σ=0.72	1.52 (1) σ=0.73	1.67 (2) σ=0.74	1.72 (2) σ=0.83	1.38 (1) σ=0.71

Table 1: Average performance relative to temperature, median in brackets where appropriate.

With respect to the other variables, intensity of change was found to have a significant effect on the number of thresholds produced (χ^2 (2)=192.412, p <0.001). The number of stimuli detected increased with intensity, with means of 53.55% (σ=49.93), 83.93% (σ=36.77) and 90.26% (σ=29.68) for 1°C, 3°C and 6°C intensities respectively. *Post hoc* Wilcoxon *T* comparisons showed significant differences in the number of detections between all stimulus intensities: 1°C vs. 3°C (z=9.62, T = 16884.0, p<0.001), 1°C vs. 6°C (z=11.84, T=19458.0, p<0.001) and 3°C vs. 6°C (z=3.08, T=2028.0, p<0.002). There was no significant effect of direction of stimulation. ROC had a significant effect (z=2.54, T=13375.0, p=0.011) with the higher ROC having a better detection rate (78.47% σ=41.13) than the lower ROC (73.06% σ=44.39). Location had a significant effect (z=4.16, T=9545.0, p<0.001) with the Thenar (80.41% σ=39.72) having a better detection rate than the wrist (71.14% σ=45.35). All of these findings are consistent with those of previous evaluations of thermal stimuli [23].

Time to Detection

Ambient temperature was found to have a significant effect on time to detection (χ^2 (4)=28.4, p <0.001). Once again performance peaks in the 15-20°C range, see Table 1. This is further indication that thermal interfaces may be more sensitive to ambient temperature than indicated by prior research. Humidity was found to have a significant effect on the time to detection (χ^2 (6)=13.948, p =0.03). *Post hoc* Wilcoxon T comparisons showed significant effects at a number of values; however there was no pattern in the differences between humidities.

		<10°C v 10-15°*C*	*<10°C* v 15-20°*C*	*<10°C* v 20-25°*C*	*<10°C* v >25°*C*	10-15°*C* v 15-20°*C*	10-15°*C* v 15-20°*C*	10-15°*C* v >25°*C*	15-20°*C* v 15-20°*C*	15-20°*C* v >25°*C*	15-20°*C* v >25°*C*
Number of Detections	Z	1.616	0.928	0.686	5.632	3.533	1.483	5.000	1.859	6.040	5.374
	p	0.106	0.353	0.493	0.000*	0.000*	0.138	0.000*	0.063	0.000*	0.000*
Time to Detection	Z	2.307	5.867	1.706	0.860	3.713	3.400	0.037	6.456	3.291	1.683
	p	0.021*	0.000*	0.088	0.390	0.000*	0.001*	0.970	0.000*	0.001*	0.092
Perceived Comfort	Z	3.542	6.597	2.024	2.097	2.365	1.84	2.583	6.074	2.462	2.495
	p	0.001 *	0.000 *	0.043*	0.036*	0.018*	0.066	0.010*	0.000*	0.014*	0.013*

Table 2: Pairwise comparison of different temperature bands using Wilcoxon Sign Ranked test. Intensity not shown as it had no significant effect.

	30%	*40%*	*50%*	*60%*	*70%*	*80%*	*90%*
Detection Rate (Accuracy %)	57.41, σ=49.60	77.45, σ=41.99	81.44 σ=38.94	81.22 σ=39.11	79.09 σ=40.76	82.11 σ=38.53	34.37 σ=48.26
Detection Time (Sec.)	2.91(2.59), σ=1.50	2.99(2.6) σ=1.53	3.69 (3.31) σ=2.02	3.06(2.61) σ=1.57	3.22(3.02) σ=1.41	3.26(2.69) σ=1.61	4.99(4.85) σ=2.53
Comfort (Scale 0-6)	2.30 (2), σ=1.39	1.82 (2) σ=1.57	2.67 (2) σ=1.55	2.42 (2) σ=1.53	1.74 (1) σ=1.39	1.27 (1) σ=1.59	1.63 (1) σ=1.36
Intensity (Scale 0-3)	1.53 (1), σ=0.72	1.67 (1.5) σ=0.76	1.61 (2) σ=0.81	1.47 (1) σ= 0.72	1.69 (2) σ=0.75	1.71 (2) σ=0.79	1.60 (1.5) σ=0.69

Table 3: Average performance relative to humidity, median in brackets where appropriate.

Intensity of change was found to have a significant effect on time to detection (χ^2 (2)=56.167, p<0.001). The time to detection decreased as the intensity increased with means/medians of 3.98/3.59 (σ=1.97) seconds, 3.04/2.64 (σ=1.59) seconds and 3/2.74 (σ=1.49) seconds for 1°C, 3°C and 6°C respectively. *Post hoc* Wilcoxon *T* comparisons showed significant differences in the time to detection between 1°C and the other intensities: 1°C vs. 3°C (z=6.22, T=22642.0, p<0.001) and 1°C vs. 6°C (z=6.58, T=22873.0, p<0.001). Direction of stimulation had a significant effect on time to detection (z=5.71, T=251982.0, p<0.001) with cool stimuli (mean=3.04, median=2.61, σ=1.69) being detected more quickly than warm stimuli (mean=3.38, median=3.01, σ=1.67). Location had an effect on time to detection (z=5.47, T=76690.5, p<0.001). With stimuli on the Thenar (mean=3.04, median=2.59, σ=1.69) detected significantly more quickly than on the wrist (mean=3.48, median=3.1, σ=1.68). ROC (z=5.37, T=45071.5, p<0.001) also had an effect on time to detection with the higher rate of change resulting in a faster time to detection (mean=3.01, median=2.64, σ=1.59) that the lower rate (mean=3.51, median=3.08, σ=1.78). Once again these results are consistent with the results of previous evaluations [23].

Subjective Stimulus Comfort

Ambient temperature was found to have a significant effect on perceived comfort (χ^2 (4)=34.247, p<0.001). Once again, the best comfort ratings occur in the 15-20°C range indicating the fine grained nature of the impact of ambient temperature (see Table 1 and Table 2). Humidity had a significant effect on perceived intensity (χ^2 (6)=21.338, p=0.002). Once again as with time to detection there is no real pattern in the differences. Intensity of change had a significant effect on perceived comfort (χ^2 (2)=35.154, p<0.001). *Post hoc* Wilcoxon *T* comparisons showed significant differences in the time to detection between 6°C and the other intensities: 1°C vs. 6°C (z=4.28, T=11936.0, p<0.001) and 3°C vs. 6°C (z=5.39, T=28537.0, p<0.001). The perceived comfort decreased as intensity increased with means/medians of 1.73/1 σ=1.37, 1.89/1 σ=1.45 and 2.58/2 σ=1.71 for 1°C, 3°C and 6°C intensities respectively. Direction had a significant effect on perceived comfort (z= 4.21, T=148932.0, p<0.001) with cool stimuli (mean=1.98, median=2, σ=1.47) being more comfortable than warm (mean=2.28, median=3, σ=1.68). Location had an effect on perceived comfort (z=2.16, T=38112.5, p=0.031). With stimuli to the Thenar (mean=2.1, median= 2, σ=1.59) being more comfortable than to the wrist (mean=2.16, median=2, σ=1.56). ROC (z=4.47, T=41652.0, p<0.001) also had an effect on comfort with the higher rate of change (mean=2.37 median =2, σ=1.67) being less comfortable than the lower rate of change (mean=1.86 median = 1, σ=1.43). As with detection rate and time to detect these findings are consistent with the results of evaluations conducted by Wilson et al. [23].

Subjective Stimulus Intensity

Statistical tests showed that ambient temperature did not have an effect on perceived intensity, as can be seen in Table 1 the figures for perceived intensity are in a similar range. Also humidity did not have a significant effect on intensity, as can be seen un Table 3.

As expected, intensity of change had a significant effect on the perceived intensity (χ^2 (2)=174.69, p<0.001). *Post hoc* Wilcoxon *T* comparisons showed significant differences in the perceived intensity of change between all intensities: 1°C vs. 3°C (z=7.53, T=6342.0, p<0.001), 1°C vs. 6°C (z=10.78, T=14406.5, p<0.001) and 3°C vs. 6°C (z=9.51, T=25474.0, p<0.001). The perceived intensity increased as the intensity increased with means/medians of 1.09/1 σ=0.49, 1.43/1 σ= 0.61 and 1.99/2 σ= 0.78 for 1°C, 3°C and 6°C intensities. Direction had a significant effect on perceived intensity (z=3.582, T=113226.5, p<0.001) with cool stimuli (mean=1.45, median=1, σ=0.79) perceived as less intense than warm (mean=1.57, median=1, σ=0.81). Location did not have an effect on perceived intensity. ROC also had an effect on intensity (z=7.02, T=33958.5, p<0.001) with the higher rate of change (mean=1.74, median =2, σ=0.81) being more intense than the lower rate of change (mean=1.39, median = 1, σ=0.64).

Direction of Change in Relation to Ambient Temperature

To provide a more in depth analysis of the effect of direction of stimulation and ambient temperature, we compared warm and cool stimuli in all temperature bands using Wilcoxon sign ranked tests. There was no effect of direction on detection rate in any band. For detection time there was an effect in the most sensitive temperature bands 10-15°C (z=2.25, T=10456.0, p=0.025) and 15-20°C (z=4.53, T=16293.0, p<0.001). For comfort only the 10-15°C (z=3.18, T=2889.5, p=0.001) had an effect with cool being more comfortable than warm. Finally for intensity there was an effect for the most sensitive temperature bands 10-15°C (z=2.58, T=4271.0, p=0.01) and 15-20°C (z=1.98, T=6339.5, p=0.048), where cool was less intense than warm. These results are interesting, warming or cooling changes were not affected by more extreme warm or cool conditions, but rather by the more favourable ambient conditions. In future work we plan to investigate this in more detail with larger temperature ranges.

Repeated Sessions

Between Session Comparison

	Session1	Session2	Session3	Session4	Session5
Detection Rate (Accuracy %)	87.03 σ=33.67	84.65 σ=36.12	81.14 σ=39.21	81.36 σ=39.03	87.61 σ=33.02
Detection Time (Sec.)	2.69 σ=1.54	2.94 σ=1.83	2.61 σ=1.38	2.77 σ=1.51	2.43 σ=1.36
Comfort (Scale 0-6)	2.42 σ=1.47	2.65 σ=1.27	2.59 σ=1.34	2.50 σ=1.35	2.69 σ=1.27
Intensity (Scale 0-3)	1.49 σ=0.69	1.55 σ=0.73	1.41 σ=0.67	1.41 σ=0.72	1.42 σ=0.67
Temperature Avg/Min/Max	11.31/ 9.25/ 14.15	16.01/ 12.40/ 20.5	16.76/ 10.35/ 24.60	16.75/ 14.55/ 18.90	22.02/ 16.50/ 27.75
Humidity Avg/Min/Max	65.68/ 56.2/ 78.95	50.95/ 44.85/ 58.90	68.59/ 51.60/ 79.10	50.86/ 45.60/ 59.50	51.55/ 33.10/ 87.70

Table 4: Average performance relative to temperature, median in brackets where appropriate.

As well as analysing the performance of the single session participants, the performance and responses of the repeated users were analysed relative to each other and relative to the single session participants. All sessions were compared using Friedman's Analysis of Variance by Ranks, with individual sessions compared using Wilcoxon pair-wise comparisons. The different sessions did not have an effect on detection rate (χ^2 (4)=6.5, p=0.165). The time to detection was significantly different between the sessions (χ^2 (4)=12.94, p<0.012). Details of the average detection times are available in Table 4, with Session 2 having slowest detection time and Section 5 the fastest detection time. Overall there were significant differences between the following sessions a number of sessions, but there was no pattern in the differences. There was no effect of repeated sessions on comfort (χ^2 (4)=8.665, p=0.07). There was no effect of repeated sessions on intensity (χ^2 (4)=6.452, p=0.168). These results indicate that training did not improve or degrade user performance. However, to ensure larger changes in environmental conditions we used a large time gap between sessions (1 month), perhaps in other conditions with shorter gaps between sessions a training effect could be found.

Comparison to Single Users

To provide a more in depth analysis of any training effects, comparisons were made between the repeated user and single user sessions based on temperature and humidity. Comparisons were made between individual sessions and the performance of the single session users for both the same temperatures and humidities using the Mann Whitney U test. While some differences were found, these did not follow any particular pattern. Table 5 provides an example of a comparison for a repeat user R2 and the general population. As can be seen, R2 has some sessions where he outperformed the general population and some where he was in line. His five sessions were spread across two temperature bands and performance was consistent within those bands. For all repeat users, there was no evidence of performance improving with more training and sessions. A full discussion of the implications of this finding and indeed all of the findings in this section is given in the next section.

	Session1	Session2	Session3	Session4	Session5
Temperature Range	*10-15*	*10-15*	*10-15*	*15-20*	*15-20*
Detection Rate (Accuracy %)	89.58*	83.33	89.58*	83.33	89.36
	74.7	74.7	74.7	84.25	84.25
Detection Time (Sec)	2.35*	3.13	2.5*	2.53*	2.52*
	3.23	3.23	3.23	3.03	3.03
Comfort (Scale 0-6)	0.77*	1.11*	1.53	1.28	1.43
	2.09	2.09	2.09	1.82	1.82
Intensity (Scale 0-3)	1.3	1.27	1.35	1.23*	1.14*
	1.52	1.52	1.52	1.67	1.67

Table 5: Repeat user R2 in comparison to general population. R2 performance is on top and general performance below. Significant differences are marked with an asterisk.

DISCUSSION

RQ1: Comparison with Existing Design Recommendations

Wilson *et al.* [23] made a number of recommendations for the design of thermal feedback for interaction: (1) the Thenar eminence is the optimal location for feedback, but non-glabrous arm locations are also suitable, (2) 1°C/sec and 3°C/sec changes are both suitable and (3) warm and cold stimuli are both suitable for use, although users have a preference for cold stimuli. Despite the different temperatures and humidities encountered by the participants in this experiment, the design recommendations made by Wilson *et al.* still hold. In our experiment, for measures of number of stimuli detected, threshold time and perceived comfort the Thenar eminence outperformed the wrist showing it to be the most sensitive area, and confirming the results of Wilson *et al.* However, while the Thenar was the best location the wrist was still suitable. Similar to the results of Wilson *et al.* [23], our results indicate that both rates of change are suitable for use, each with slightly different characteristics. While 1°C/sec changes are slower and require a longer time to detect, they are more comfortable and less intense than the larger ROC. 3°C/sec changes are more likely to be detected and are much faster to detect, but are less comfortable. Finally, although both warm and cold stimuli are equally detectable and are perceived to have equal intensity, cold stimuli are faster to detect and are more comfortable as they feel less intense. More in depth investigation of direction of change in different ambient temperatures also indicates that cool changes are preferable to warm changes, with both still being usable. Research by Halvey *et al.* [8] indicted that warm stimuli might be more comfortable, however in this evaluation different clothing materials were placed between the Peltiers and the skin. So, for direct contact with the skin cool stimuli are more comfortable based on both the results of this experiment and the results of Wilson *et al.*

RQ2: Effect of Ambient Temperature

Hirosawa *et al.* [10] investigated the relationship between room temperature and warm and cold thresholds on the middle fingertips. They found that the warm and cold thresholds were affected by room temperature changes, with a neutral zone of 15-25°C for thresholds. However, the results of our analysis indicate that the effect of ambient temperature on thermal feedback perception appears to be finer grained than previously reported. Ambient temperature had a significant effect on the number of stimuli detected, threshold time and perceived comfort. Ambient temperature was not found to have an effect on perceived intensity. Our results indicate that the optimal performance in terms of detection rate (84.25%) and time to detection (3.03 seconds) is in the range 15-20°C. With performance degrading in both warm and cool direction around the 15-20°C temperature zone. While the differences are significant, with the exception of detection rate for >25°C, all of the other results are useable and indicate that thermal interfaces could be used. It should be noted that the >25°C is one of the more sparse data points in our data and the low detection rate might be due to individual thermal sensitivity of the users, this point should be investigated further in future work. Interestingly, the 15-20°C range also had the most favourable comfort ratings; again the ratings indicated that stimuli were perceived as being less comfortable in both the warm and cool directions. It should be noted that the average comfort rating (2.13) is not in the uncomfortable range. This indicates, once again, that while ambient temperature affects comfort, the thermal stimuli used in this evaluation were still comfortable and useable.

RQ3: Effect of Ambient Humidity

Givoni *et al.* [4] reported the effect of humidity on perceived comfort to be small. However, the results of our study indicate that humidity had a significant effect on the number of stimuli detected, threshold time and perceived comfort. Pairwise comparisons for detection rate showed that the differences in detection rate were for more extreme humidities, i.e. <30% (57.41% accuracy) and >90% (34.37% accuracy), with all other humidity ranges having detection rates in the range 77%-82%. This result is in keeping with the findings of Givoni *et al.* [4] where there are some effects at extreme humidities but otherwise the differences are negligible. While significant differences were reported for threshold time and perceived comfort, the differences between humidity ranges did not follow any particular pattern. Perhaps these differences could be attributed to some sparsity for particular humidity ranges and as such those ranges were affected more by the thermal sensitivities of the individual users. Further experimentation is needed to explore this.

RQ4: Training Effects

The results from the repeated user sessions indicate that there is very little difference between performance and perception in the different sessions, as can be seen in Table 4. There was no significant difference between the five sessions for the number of stimuli detected, the perceived comfort of the stimuli and the perceived intensity of the stimuli. There was a significant effect for threshold time. However, there was no clear pattern of improvement or degradation as the exposure to the stimuli increased over the sessions, although Session 5 did have the fastest detection time (2.43 sec.) and Session 2 the slowest (2.94 sec.). To gain further insight into any possible learning effect, individual performances were compared to the general population (as represented by the single session users) for the same temperature and humidity ranges. This comparison, while showing some differences, did not reveal any patterns and certainly no increase in performance for the repeated users. These results suggest that users learn the thermal feedback quickly and it would be usable for single or repeated use in an interface.

CONCLUSIONS

This paper has presented a detailed study investigating how well users were able to detect warm and cool stimuli presented to two body locations while outdoors. Our aim was to identify the sensitivity of thermal stimulus presentation to environmental factors in order to inform thermal feedback design for HCI and as such our novel contributions can be summarised as follows 1) Proposed design recommendations for thermal interfaces are based on a very recent study [23]. The confirmation that they hold across environmental conditions is a novel, useful and non-obvious finding. In addition the confirmation that cool stimuli directly on the skin are preferred is important, as other different studies have found that in some cases warm stimuli are preferred [8, 22]. 2) Ambient temperature affects stimuli detection rates and this effect is at a more fine grained level than presented in the psychophysical literature [10]. This knowledge is essential for designing thermal UIs. 3) Humidity does not appear to have the same effect as ambient temperature with the differences being at extreme humidities. While Givoni et al. [4] had a similar finding, they were not concerned with direct stimulation as we are. 4) No training effect was found for this evaluation; however a large gap between sessions was used.

In conclusion this paper has shown that while thermal feedback is sensitive to ambient temperature that previous recommendations are still valid, and as such we can design and use thermal feedback in a wide range of usage settings.

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. Fujita, H. and K. Nishimoto, *Lovelet: a heartwarming communication tool for intimate people by constantly conveying situation data.* Proc. CHI 2004.
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. Givoni, B., J. Khedari, N.H. Wong, H. Feriadi, and M. Noguchi, *Thermal sensation responses in hot, humid climates: effect of humidity.* Building Research & Information, 2006. **34**(5): p. 496-506.
5. Gooch, D., *An Investigation into Communicating Social Presence With Thermal Devices.* MSc Dissertation, University of Bath, 2009: p. 1-390.
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.
8. Halvey, M., Y. Vazquez-Alvarez, G. Wilson, and S. Brewster. *The Effect of Clothing on Thermal Feedback Perception*. Proc ICMI 2011: p. 217-220.
9. 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.
10. 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.
11. Iwasaki, K., T. Miyaki, and J. Rakimoto, *AffectPhone: A Handset Device to Present User's Emotional State with Warmth/Coolness.* BIOSTEC2010, 2010. **2010**(B-Interface Workshop): p. 1-6.
12. 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.
13. Jones, L.A. and M. Berris, *The Psychophysics of Temperature Perception and Thermal-Interface Design.* 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002: p. 1-6.
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.
16. Kushiyama, K., T. Baba, K. Doi, and S. Sasada, *Thermo-Pict neo.* Proc SIGGRAPH 2010: p. 1-1.
17. Lee, W. and Y.K. Lim, *Thermo-Message: Exploring the Potential of Heat as a Modality of peripheral Expression.* Proc CHI Extended Abstracts 2010: p. 1-6.
18. Nakashige, M., M. Kobayashi, and Y. Suzuki, *"Hiya-Atsu" media: augmenting digital media with temperature*. Proc. CHI 2009: p. 3181-3186.
19. Narumi, T., T. Akagawa, Y.A. Seong, and M. Hirose, *Thermotaxis.* Proc. SIGGRAPH 2009: p. 1-1.
20. Stevens, J.C., *Thermal Sensibility*, in *The Psychology of Touch*, M.A. Heller and W. Schiff, Editors. 1991, Lawrence Erlbaum: New Jersey.
21. Strigo, I., F. Carli, and M. Bushnell, *Effect of ambient temperature on human pain and temperature perception.* Anesthesiology, 2000. **92**(3): p. 699-707.
22. Wettach, R., C. Behrens, A. Danielsson, and T. Ness, *A thermal information display for mobile applications.* Proc. MobileHCI 2007: p 182-185.
23. Wilson, G., M. Halvey, S.A. Brewster, and S. Hughes, A., *Some Like it Hot? Thermal Feedback for Mobile Devices*. Proc. CHI 2011: p 2555-2564.