In the Heat of the Moment: Subjective Interpretations of Thermal Feedback During Interaction

Graham Wilson[1], Gavin Davidson & Stephen Brewster[1]
Glasgow Interactive Systems Group
School of Computing Science
University of Glasgow, Glasgow G12 8QQ
[1]{first.last}@glasgow.ac.uk

ABSTRACT
Research has shown that thermal feedback can be an engaging and convincing means of conveying experimenter-predefined meanings, e.g., material properties or message types. However, thermal perception is subjective and its meaning in interaction can be ambiguous. Interface designers may not be sure how users could naïvely interpret thermal feedback during interaction. Little is also known about how users would choose thermal cues to convey their own meanings. The research in this paper tested subjective interpretations of thermal stimuli in three different scenarios: social media activity, a colleague's presence and the extent of use of digital content. Participants were also asked to assign their own thermal stimuli to personal experiences, to help us understand what kinds of stimuli people associate with different meanings. The results showed strong agreement among participants concerning what warmth (presence, activity, quality) and cool mean (absence, poor quality). Guidelines for the design of thermal feedback are presented to help others create effective thermal interfaces.

Author Keywords
Thermal feedback; interaction design; mobile interaction.

ACM Classification Keywords
H.5.2. User Interfaces – Haptic IO.

INTRODUCTION

Thermal stimulation is an inherent aspect of sensory experience, with strong links to social (e.g., physical closeness) [2,14] and emotional (e.g., "warm and loving") phenomena [20]. Thermal feedback in HCI may be capable of improving user experience by bridging the gap between data and their underlying social or emotional content. It has been used in HCI to convey information [1,21], improve materiality [5,16] or for communication [14,18]. However, most research has either been technological, developing new ways of providing thermal feedback [1], or perceptual, testing how well participants can detect or identify thermal stimuli [21,22]. There are few examples of applications where thermal feedback is tested in real-world interactions to see how users interpret it. Researchers have started to measure subjective views on potential meanings or uses for thermal feedback [3,14,18], but they have been in limited scenarios and provide few details about the specific stimuli used. Research is required to understand interpretations of thermal feedback in a range of familiar scenarios if it is to be effectively utilized in everyday interfaces.

Existing research also tends to prescribe meaning for the thermal feedback, attaching specific information to the stimuli used. Little is known about how users naturally and freely interpret thermal changes in a variety of interaction environments, which is key for the design of effective thermal UIs. This paper extends previous research by measuring the subjective meanings attributed to thermal stimuli in real-world examples. We let participants assign their own subjective meanings to stimuli and choose their own stimuli to represent personal experiences, to help us understand how people would interpret and use thermal feedback during interaction. We also tested different data types, including categories, range data and experiences, to provide a broader understanding of how thermal feedback is understood. The paper makes the following novel contributions: 1) Testing four real-world interactions not yet investigated with thermal feedback; 2) Testing associations of thermal feedback to different information/data types; 3) Recording how participants inherently interpret and assign thermal feedback; 4) Providing clear design guidelines.

We chose four scenarios to test interpretations of thermal cues: 1) conveying social media activity, 2) conveying physical presence, 3) conveying application usage and 4) a restaurant experience scenario. Thermal feedback has strong inherent emotional and social cues and so it may be useful for enhancing uses where these cues are central. Therefore, these scenarios were chosen because they are common and familiar, and they are related to social/emotional experiences, so allowing us to measure the corresponding associations of thermal feedback.

RELATED RESEARCH

Thermal feedback was first used in HCI to improve the materiality of objects in virtual reality, by mimicking different patterns of thermal conductivity [6,7]. Since then, the most common use case has been to augment media or communication with emotional or social content. There is evidence of an inherent link between biological temperature and social emotion, as Williams and Bargh [20] found a connection between physical warmth and interpersonal warmth.

Nakashige *et al.* [15] combined images of food with warm, neutral or cold stimuli and found that foods presented with the correct corresponding temperature (e.g., warmth with soup) were rated as more delicious than those that were not. Both Salminen *et al.* [17] and Halvey *et al.* [3] measured participants' emotional responses to thermal stimuli. Salminen *et al.* looked at the stimuli in isolation, while Halvey *et al.* studied the effect of combining stimuli with audio and visual media. Both found that thermal stimuli influenced emotional state but the results were slightly different. Halvey *et al.* found warm stimuli were more pleasant than cool, while Salminen *et al.* found no difference in pleasantness. Warm stimuli did generally lead to higher arousal in both studies, compared to cool. Wilson *et al.* [23] also reported that warm stimuli were more intense and less comfortable than neutral and cold stimuli. These emotional studies did not measure any perceived meaning in the stimuli, only the participants' resulting affective state.

Researchers have looked at using thermal stimulation in interpersonal communication. Iwasaki *et al.* [8] augmented a mobile device with galvanic skin response (GSR) sensors to convey the emotional state of another user, with higher emotional arousal resulting in warmer feedback. Gooch [2] also showed that providing warm stimuli around the abdomen (to mimic a "hug") during instant messaging between physically separate users could increase feelings of social presence, although the effect was quite weak.

These studies all presented participants with the researchers' own choice of stimuli, sometimes with a prescribed meaning, before measuring participant responses to them. But how would participants naturally interpret meaning from ambiguous thermal changes, and how would they convey their own intentions through thermal feedback? Understanding this is key if thermal displays are to be effective on a large scale. Lee and Lim [13,14] have investigated participants' own subjective perceptions of thermal sensations in general [13] and specifically in the context of interpersonal communication [14]. In the latter study, they also asked participants to design their own feedback choices along three dimensions: temperature, duration and rate of temperature change. Participants tended to treat the warm-cold dichotomy as two opposites of meaning, with stronger changes representing the degree of difference.

Participants stated that there were specific temperatures that were appropriate for particular phenomena, suggesting a believed universality in interpretation. However, the results did not necessarily support this. In general, warmth was used for positive meanings, such as physical attraction or enjoyment, while cold represented negative meaning, such as the presence of a stranger. However, some participants used cold to represent positive aspects, such as refreshment.

Suhonen *et al.* [18] tasked pairs of participants with discussing something happy, something sad or angry and something emotionally neutral (restaurants). They allowed users to send warm or cold stimuli based on their own intentions and recorded what meaning they attached to the stimuli. Like Lee and Lim [14], warm sensations were generally used to convey positive and pleasant feelings or experiences (in both the happy and sad/angry scenarios) but across a wide range of interpretations, including emphasising happy memories, social closeness, empathy, gratitude and good food/restaurants. However, some participants used heat to indicate anger or annoyance. In contrast, cold was regularly associated with negative factors, such as nervousness, sadness, pain or anger. Cold represented a poor choice of restaurant. The meaning of stimuli, and the valence attached, depended on the valence of the context (discussion topic), which is in line with Lee and Lim's [14] suggestion that emotional state influences interpretation.

These papers are important as they give some insight into how naïve participants would use and interpret thermal feedback in the real world, outside of prescribed experimental stimuli. However, there are several limitations. Neither Lee and Lim [14] nor Suhonen *et al.* [18] report on the specific thermal feedback designs used in their studies, in terms of temperatures, rates of change or durations, so it is not known how the participants' intentions map to specific thermal stimuli. This information is needed to design appropriate feedback in the future. Also, the feedback designs and interpretations are limited to only interpersonal communication. In this paper we present an investigation into how interpretations vary across different scenarios and in relation to different subject matter, and outline specific thermal feedback design guidelines based on the results.

EXPERIMENTS

This section describes the four experimental scenarios designed to measure the subjective interpretation of thermal feedback: social media activity, physical presence, content deletion and restaurant experience. Like Lee and Lim [14], we did not dictate specific mappings of feedback to meaning. In each example, we present a range of thermal stimuli and ask participants what meaning or information they take from the stimulus. 15 participants (3 F) aged 18 to 31 (mean = 22.7) took part in all scenarios in a random order and were paid £6 for a 60min session. *A priori* sample size computation indicated 15 was sufficient for valid analysis.

Thermal Apparatus

The thermal stimulation was provided by the Peltier-based device used by Wilson *et al.* [21,22] (Figure 1). It is controlled over Bluetooth and can be set between -20°C and 45°C, accurate to 0.1°C. We used two 2cm^2 Peltier modules and changed temperatures at a rate of 3°C/sec to maximise the sensation [23]. For all four scenarios, the Peltiers were sitting on a desk facing up for the participants to rest the palm of their hand on top, supported by a padded rest (Figure 2). The Peltiers were controlled by either a PC (in the restaurant scenario) or a mobile phone (other scenarios). While we chose to stimulate the palm, arm locations are similarly sensitive [23], so wearable devices, such as smart watches, may be suitable stimulators for mobile interaction. Mobility influences thermal perception [21,22,23] so this will be tested in future research.

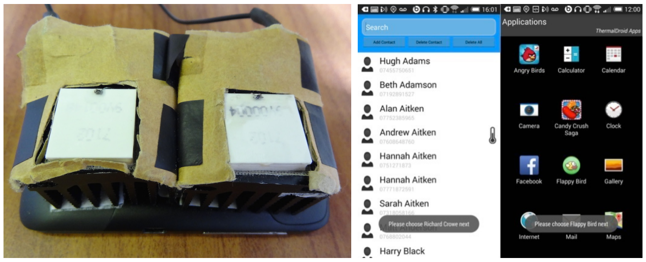

Figure 1: Peltier devices (left) and Contact List/Content Deletion interfaces (right). Cardboard covered the heatsinks.

Lee and Lim [14] allowed their participants to vary three parameters of thermal change: temperature delta (*t*Δ, the change from skin temperature), rate of temperature change (ROC) and duration. Unfortunately, they did not report what range of ROC or durations were used, nor what specific designs participants chose. Perceptual research suggests that increasing any one of these factors would cause an increase in the intensity of the sensation [10,23]. Further, immediate perception of thermal stimuli is not guaranteed [4,11,23] so the duration and ROC may not be reliably perceivable in realistic scenarios [4,22,23]. We chose to limit our stimuli to changes in *t*Δ, which will influence intensity but will not require the accurate perception of other factors.

Figure 2: Experimental setup with Peltier elements under the palm and the padded armrest for comfort.

Experimental Setup

Thermal perception is different to audio, visual and tactile perception in that it is bipolar: the skin rests at a homeostatic neutral temperature and can be warmed or cooled from there [9]. Other modalities are unipolar, changing from no stimulus to increasing levels of stimulus. A resting state (in between trials) for audio, visual and tactile stimuli involves the absence of a stimulus. This is not possible for thermal feedback, as the stimulator and skin always have a temperature, and so the resting state for thermal feedback is skin temperature. For all interactions, the Peltiers were returned to a neutral temperature of 30°C between each trial: thermal research commonly uses a similar set starting temperature for controlled comparison between stimuli and 30°C is within the skin's natural range of resting temperatures [9]. In the social media, presence/availability and content deletion scenarios, 30°C was also included as an experimental stimulus (as it is a valid potential interaction cue), so participants were made aware that any lack of change during a trial was intentional and they were to treat the stimulus like any other. The specific temperatures used in each scenario are described below, but they all ranged from 22°C to 38°C. This range was chosen because it is safe, comfortable, reliably perceivable and centered on neutral 30°C skin temperature [9,12,23]. These temperatures have also elicited emotional responses in previous research [3,17,18].

For all four scenarios, the participant was sat at a desk in an office. On the desk were the armrest, Peltier devices, an Android mobile phone and a computer monitor and mouse. The participants rested their non-dominant hand on the armrest so that the palm of their hand made good contact with the two Peltier modules. The hand remained in contact with the Peltiers throughout each scenario but was removed during rest periods between scenarios. For the social media, physical availability and content deletion scenarios, the participant interacted with the mobile phone (which controlled the Peltiers) with their dominant hand to receive experimental instructions and provide input via the touchscreen GUI. For the restaurant experience scenario, the participant used the mouse to interact with an interface shown on the monitor while a PC controlled the Peltiers. A PC was used because the task required participants to type text and search and/or scroll long lists to find bars and restaurants, and they were presented with images and text describing the establishments. This interface would have been more cluttered and cumbersome on a small screen.

Online Activity: Phone Contact List

This scenario investigated the use of thermal feedback to provide an immediate overview of online activity, by conveying the recency of an individual's social media activity from their entry in a contact list application. The purpose was to see how participants relate temperature to temporal activity. As the contact list is scrolled, the user can hold a finger over an individual's contact to receive thermal feedback relating to how recently the person has posted on social media. We presented a range of temperatures and asked participants to state how recently the person was active.

Stimuli & Measures

Nine thermal stimuli were used: 22°C to 38°C in 2°C intervals. 1°C changes can be difficult to perceive [23], so a 2°C change was chosen as the smallest usable delta. These nine temperatures were each assigned to two different contact names and all 18 names were used as targets for the experimental trials. This meant that each stimulus was responded to twice but the participant was not necessarily aware of it, reducing any bias. For each trial, the Peltiers would be changed to one of the temperatures and remain there until a response was submitted. The Peltiers were then returned to 30°C for 10 seconds in preparation for the next trial. Participants were told that the thermal feedback represented how long ago the person was active on social media, but not how the temperature related to time. Once the target name and accompanying temperature were presented, the user was asked the question *"how long ago was this person on social media?"* A text box for a number value and a drop-down menu to indicate the time frame were presented: "seconds", "minutes", "hours", "days", "weeks", "months" or "years". This scenario used categorical data representing seven time frames to investigate how consistently thermal stimuli were attributable to 1) a set range of categories and 2) time.

Procedure

The interface (Figure 1, right) presented the participant with an alphabetical contact list of names and the experimental software requested each of the 18 target names in a random order, one per trial. In a given trial, the software would present the text *"Please choose [name] next"*. The participant would then have to scroll the list to find the name and long-press it (touch for 1 second), at which point the Peltiers would change temperature from neutral to the accompanying temperature. The interface then presented the experimental question along with the text box and drop-down menu for user input. Once the user had responded, he/she pressed a "submit" button, after which the Peltiers were returned to 30°C for 10 seconds before the next trial began.

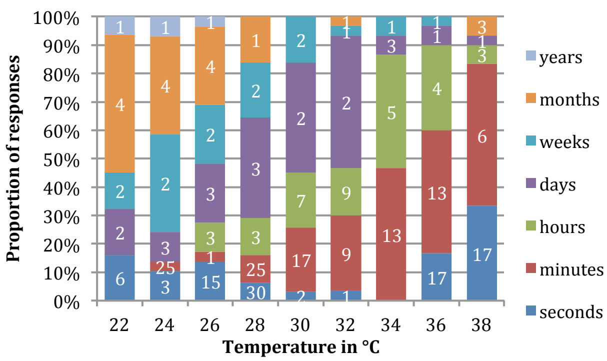

Figure 3: Bars show the proportion of responses that attributed each time frame to each thermal stimulus in the Contact List scenario. The numbers within each bar show the average enumeration of that time frame, e.g., 17 seconds.

Results & Initial Discussion

Figure 3 shows the proportion of responses that attributed each time frame to each of the nine thermal stimuli. These data do not take into account the number value enumerating the time frame, such as the 3 in "3 weeks", but the data were organized so that responses such as "14 days" were counted as a response in "weeks" rather than "days". The results show that participants generally attributed colder stimuli to longer time frames (older activity), such as days, weeks and months, and attributed warmer stimuli to shorter time frames (more recent activity) such as minutes or hours. The average value enumerating each time frame (the 3 in "3 weeks") is shown in white within each bar in Figure 3. The cooler stimuli were perceived as representing between 3 seconds and 1 year ago, while the warmer temperatures were interpreted as being between 17 seconds and 3 months ago. Overall, the colder the feedback temperature, the older the social media activity is perceived as being.

As seen in Figure 3, ~90% of responses interpreted the warm temperatures (34-38°C) as representing activity ≤ 1 day ago. For the warmest temperatures (36°C and 38°C), most were interpreted as being just seconds or minutes ago. There is a sudden shift where the neutral temperatures (28-32°C) are mostly interpreted as representing days ago. This trend of older activity continues as the majority of responses interpreted colder temperatures (22-26°C) as meaning weeks, months and even years since the last activity. To identify which temperature was most associated with each time frame, Friedman tests were run on the number of responses for each temperature within each time frame, e.g., compare the number of times "seconds" were attributed to each temperature. Following a significant Friedman test, *post hoc* Wilcoxon tests with Bonferroni adjusted p-values ($p<0.0014$) were used for pairwise comparisons. Any significant differences would indicate the most associated temperature(s). Effect size was calculated on the Wilcoxon tests using $r = Z/\sqrt{N}$, where Z is the Wilcoxon Z statistic and N is the total number of samples. Effect sizes were interpreted as $r = 0.1$ (small), $r = 0.3$ (medium) and $r = 0.5$ (large). A summary of the statistical analysis, including recommended feedback design, is shown in Table 1.

Number of responses attributing each temperature to time frames				
Time Frame	**Friedman Test (n) = df**	**Significant Pairwise Comparisons**	***Best Conveyed By***	*Suggested Alternative*
Seconds	$\chi^2(8)=43.10$, $p<0.001$	None	~	*38°C*
Mins	$\chi^2(8)=80.69$, $p<0.001$	34°C, 36°C, 38°C > 22-28°C	***> 34°C***	~
Hours	$\chi^2(8)=54.31$, $p<0.001$	34°C > 22°C, 24°C	***34°C***	~
Days	$\chi^2(8)=66.95$, $p<0.001$	32°C > 24°C	***32°C***	~
Weeks	$\chi^2(8)=43.89$, $p<0.001$	None	~	*24°C*
Months	$\chi^2(8)=76.0$, $p<0.001$	22°C > 30-38°C	***22°C***	~
Years	Not significant	None	~	*22°C*

Table 1: Summary of statistical analysis on the Contact List data. The table includes the "best" temperature for conveying each time frame and suggested alternative temperatures.

For the "seconds" data, there was a significant effect of temperature on the number of responses, but no pairwise comparisons reached the adjusted p-value, suggesting no individual temperature can reliably convey seconds. There was a significant effect within the "minutes" data with several significant pairwise comparisons: all three of 22°C, 24°C and 26°C had significantly fewer "minute" responses than 34°C, 36°C and 38°C (all medium effect sizes: $r =$

0.31 to 0.35). 28°C had significantly fewer than 34°C (r = 0.3) and 38°C (r = 0.31). This shows that warmer temperatures (≥34°C) are much more associated with a minute time frame than cooler temperatures, but no singular warm temperature is best suited to representing minutes. A significant effect was found within the "hours" data, with the 34°C stimulus having significantly more "hour" responses than 22°C (r = 0.38) and 24°C (r = 0.38). As 34°C was the only value significantly higher, it might best represent hours.

Following a significant effect within the "days" data, only 32°C had significantly more responses than 24°C (r = 0.31), suggesting 32°C is good for representing days. While there was a significant effect in the "weeks" data, no pairwise comparisons reached significance, suggesting no individual temperature conveys weeks reliably. A significant effect was found within the "months" data, as 22°C had significantly more responses than 30-38°C (r = 0.40 to 0.42), suggesting it is best at representing months. Finally, there was no significant effect of temperature on the "years" data.

Overall there is a greater sense of recency attributed to warmer stimuli. A pattern that goes against this trend is the somewhat U-shaped relationship between temperature and the number of "seconds" responses. The number decreases from 38°C (33% of responses) to 34°C (0%) before increasing again, with 16% of responses interpreting the coldest temperature (22°C) as representing seconds. Therefore, responses for 22°C ranged from seconds to years. Similarly, 6% of 38°C responses were months, giving a range of seconds to months for the hottest temperature. It may be that, for some participants, extreme temperatures can be interpreted as either extreme value within the relevant range.

Physical Presence: Augmented Office Door Handle

This scenario investigated how participants relate temperature to the presence and availability of a colleague in an office environment. Here we envision a smart office with an augmented door handle capable of warming up and cooling down (Figure 4). An individual wants to speak with a coworker but is unsure of his/her availability. If the inside of the office cannot be seen, the physical presence of the person is unknown. An un-answered knock on the door could mean that 1) they are away, 2) they are in but do not want to be disturbed or 3) they simply did not hear the knock.

For this scenario, we imagined conveying the co-worker's presence through thermal feedback when the visitor touched the augmented door handle. We chose a set of five 'availability' categories that cover a range of situations: 1) "Out of department", 2) "Back soon", 3) "In. Please knock", 4) "Available for short times" and 5) "Extremely busy". These were modelled on the paper indicators commonly used on office doors where a marker indicates the appropriate category of busyness. During the study, participants felt different stimuli and had to interpret their meaning as one of the five categories. While we prescribed the categories, we did not attach specific stimuli to them: participants applied their own interpretation to each stimulus.

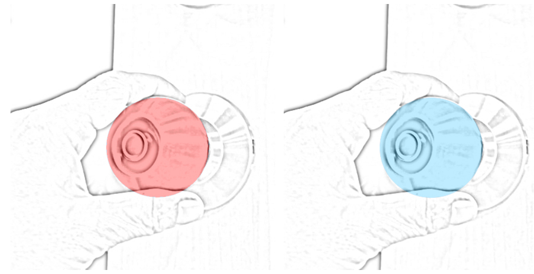

Figure 4: An augmented door handle could use thermal feedback to convey the presence or absence of the person inside.

Stimuli & Measures

Five stimuli were used: 22°C, 26°C, 30°C, 34°C and 38°C and each was presented twice, one per trial and participants were told that the thermal feedback represented the person's "availability". After the Peltier temperature had been changed, the mobile device presented a screen with the request: *"Choose which option best suited that temperature?"* There was a drop-down menu showing the five availability categories for participant responses. This scenario also used categorical data, but to test how thermal feedback is mapped to the physical presence of a person.

Procedure

During each trial, the software would set the Peltiers to the random target temperature before presenting the request to choose the corresponding meaning. Participants chose the availability category they felt was best represented by the thermal stimulus before pressing "submit". The Peltiers returned to 30°C for 10 seconds before the next trial began.

	Absence		*Presence*		
°C	**Out of Dept.**	**Back soon**	**In. Please knock**	**Short times**	**Extremely busy**
38	3	1	1	4	21
34	1	5	11	9	4
30	2	6	20	2	0
26	13	9	4	3	1
22	19	4	3	0	4
Best	<26°C	26°C	30°C	34°C	38°C

Table 2: Frequency counts for each label attributed to each thermal stimulus during the Physical Presence scenario. Shading shows most common (green) to least common (white). The "best" temperature for conveying each state is also shown.

Results & Initial Discussion

The results showed a strong degree of agreement among participants as to how temperature related to the presence of the colleague, which is inherently linked to his/her availability (see Table 2). In accord with previous research [2,14], the warm temperatures were interpreted as conveying the presence of the colleague but with greater warmth indicating less availability (while present). In contrast, cool temperatures were interpreted as absence, with greater cold indicating a greater degree of unavailability (absent for longer). The results for the moderate warm (34°C) are mostly split between "In. Please knock" (37%) and "Available for short times" (30%). There is a similarity in the practical meaning of these two categories, as both could mean that the visitor is welcome to enter. However, 67% of the responses for the neutral 30°C were attributed to "In. Please knock", suggesting that the extra warmth from 34°C led some to interpret it as representing less availability.

The same analysis as for the Contact List task was employed: comparing the number of responses for each temperature within each availability category, to identify which temperature was particularly associated with it. Significant Friedman tests were followed by Bonferroni-adjusted ($p<0.005$) Wilcoxon pairwise comparisons and effect size r calculations. There was a significant effect of temperature on the number of responses under all five categories ($p <= 0.001$). For the "Out of Department" responses, both 22°C and 26°C had significantly more responses than 30°C, 34°C and 38°C ($r = 0.36$ to 0.49), showing that the coolest temperatures were uniquely associated with the person being away. For the "Back Soon" data, 26°C had significantly more responses than 38°C ($r = 0.4$), suggesting that cool temperatures are more associated than particularly warm ones. The "In. Please Knock" data showed that 30°C had significantly more responses than all other temperatures ($r = 0.34$ to 0.49), and 34°C had significantly more than 38°C ($r = 0.36$) and 22°C ($r = 0.32$). Moderate temperatures are, therefore, particularly associated with presence and availability. For "Short times" 34°C had significantly more responses than 22°C (large effect size: $r = 0.5$), suggesting moderate warmth represents this state of moderately busy presence. Finally, 38°C had significantly more responses than all other temperatures under "Extremely Busy" (large effect sizes: $r = 0.53$ to 0.59), showing that the warmest temperature uniquely represented the highest busyness.

Ordinal data: Content Deletion

Temperature has a strong association with safety and danger in our environment: burning fire, comfortable warmth and freezing cold. For this scenario we wanted to investigate how thermal feedback might be interpreted in a situation that related to dangerous or risky interaction. There are a number of computing functions that result in permanent and irrevocable changes to the system, such as deleting content, upgrading an operating system or reverting to a factory state. Operating systems often force the user to reflect on the potential severity of the action and confirm their desire to carry it out through a dialogue box. However, the severity is likely to depend on the significance of the content being altered. Deletions, for example, may not matter if the content in question is unused or unwanted. If it is frequently used, the deletion is potentially more dangerous.

In this scenario, we looked at how thermal feedback might relate to the severity of deletions in the context of removing applications from a mobile phone, by using the feedback to convey the amount of use the application has had (although the principle could be applied to any digital content). When the participants long-pressed an icon (the action used in Android and iOS to allow deletions) they were provided with thermal feedback indicating the amount of use of the associated application. Different interactions require different data types, so a second motivation for this study was to measure interpretation of thermal feedback along an ordinal scale (in this case magnitude of usage), as the previous scenarios measured allocation of stimuli into categories.

Stimuli & Measures

The same nine stimuli from the Contact List scenario were used again. Each temperature was assigned to two of 18 application icons, each of which was a target during the experiment. Like the Contact List, this meant that each temperature was tested twice, but the user was not primed by a previous response. The participants were told that the thermal feedback represented how much the relevant application had been used but not how the temperature related to usage. The study followed a *magnitude estimation* design, where the participants were asked to attribute numbers to each stimulus that represented the magnitude of application usage. In perceptual research, magnitude estimation studies typically test only heat or cold perception and not both directions in one study [12]. In these cases, the stimulator rests at neutral skin temperature and changes to varying degrees of warmth *or* cold, but not both. The participant then attributes magnitude values to those varying degrees. As we wanted to examine interpretations of both within the same application, we changed the design: we made neutral indicate a moderate amount of use and allowed participants to attribute stimuli along a spectrum of greater or lesser use.

Before the experiment started, the Peltiers were set to 30°C and participants were told that it represented a middle amount of usage and was given the magnitude value of 50. After each stimulus, the user was asked *"how much has this application been used?"* Participants were instructed to respond with a value that represented their perceived magnitude of usage by typing numbers into a text box. At this point, we only had a number attributed to each temperature but we would not necessarily know whether that number represented more or less usage than the middle 50. For example, 70 could represent more usage (e.g., a greater amount) or less usage (e.g., it was used longer ago). Therefore, the application also asked participants, via a drop-down box, whether their value represented a greater or lesser extent of usage than the middle value of 50. In contrast to the first two scenarios, this used a participant-defined, continuous range. While we dictated the middle value, participants could attribute any number to the stimuli, including negative numbers and large positive numbers. This scenario allowed us to see how thermal feedback is mapped to a continuous range of data, representing an overall "amount", which could be applied in other scenarios, such as navigation distance or e-fitness targets.

Procedure

The procedure was very similar to the Contact List scenario, however the experiment GUI (Figure 1, right) showed 18 icons in an arrangement similar to both Android and iOS. During each trial, the experimental software presented the instruction *"Please select [application] next"* and users were told to long-press the icon to delete it. Each of the 18 icons was a target, one per trial. Upon long-pressing, the software changed the Peltiers to the relevant temperature and participants were presented with the text box and drop-down menu. After pressing "submit", the Peltier was returned to 30°C and the next trial began after 10 seconds.

Results & Initial Discussion

The geometric mean magnitude values for each temperature are shown in Figure 5. There was a significant positive correlation between stimulus temperature and perceived usage magnitude using Pearson's product-moment correlation coefficient (r (7) = 0.989, $p<0.01$). Perceived usage increased as stimulus temperature increased. Every participant indicated that higher values indicated a higher level of usage, which suggests that warmth is interpreted as representing high usage and cold temperature low usage. While there is a strong correlation, the pattern is slightly different for warm and cold stimuli. Warm temperatures (>30°C) increase in usage magnitude linearly, but magnitudes for cooler temperatures (<30°C) drop more quickly before levelling out at 24°C.

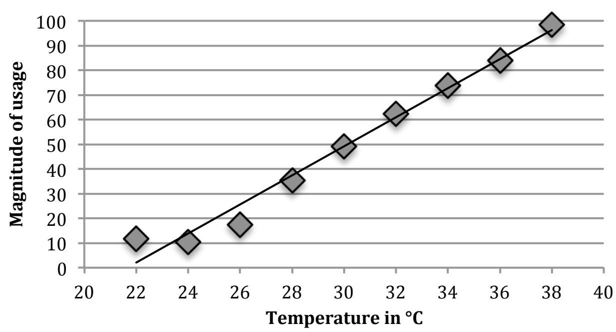

Figure 5: Geometric mean of magnitude values given for each temperature during the application deletion task (*r* = 0.989).

A repeated-measures ANOVA on normalized magnitude values showed a significant effect of temperature on usage magnitude ($F_{(6,232)} = 68.044$, $p<0.001$). *Post hoc* Bonferroni pairwise comparisons showed that 22°C, 24°C and 26°C were not significantly different from each other, but all other temperature comparisons differed significantly ($p<0.05$ for all adjacent temperatures except 32°C vs. 34°C; $p<0.001$ for all other comparisons including 32°C vs. 34°C). Effect sizes were all r = 0.3 or above. This shows that all warm temperatures had significantly different perceived magnitude values, but there was no difference between temperatures below 28°C.

It may be that participants perceived cold as representing a more singular state of "little use", while degrees of warmth meant degrees of use. While users can have some difficulty differentiating degrees of change in thermal feedback, this affects warm temperatures much more than cold [4,22,23]. This is, therefore, less likely to have affected perception of the cold stimuli and the linear, significant increases in perceived usage above 30°C would suggest that this did not happen to the warm stimuli either. While the mapping of warmth to greater usage is not particularly surprising, the finding that users can reliably appreciate up to 7 temperature levels is a new finding in thermal feedback research. From the results, it appears as though thermal feedback could effectively convey amount of usage. As mentioned above, application usage and importance are related, and so thermal feedback may also be useful in supporting users to reflect on the importance of their actions in different applications. These could include agreeing to share personal information or opening potentially suspicious emails/links.

Subjective Experience: Restaurant Ratings

In the previous examples, we presented the participants with set thermal stimuli and asked for their interpretations of those signals. Previous research has also provided set stimuli for users to interpret. Only Lee and Lim [14] have asked participants how they would convey their own information or intentions by choosing their own feedback designs, and this was limited to interpersonal communication. However, Lee and Lim provided no details about the actual stimuli participants' chose. Here we take a step further by asking participants to convey their personal opinions on familiar bars or restaurants by setting thermal stimuli to a desired temperature. We report and discuss the specific temperatures associated with each type of response.

Restaurant review sites are widely used and the numerical rating, typically from 1 (bad) to 5 (excellent), is an established means of quickly identifying "good" or "bad" places to go. However, individual reasons for liking or disliking an establishment include many social/emotional aspects, such as the quality of food, the service or the atmosphere. Temperature and food are inherently linked, e.g., soup is generally hot and soft drinks are generally cold, and thermal stimuli are appreciated in food-based media [15]. On a hedonic level, hot drinks may be enjoyable on a cold day, while cold drinks may be enjoyable on a hot day. The link between language and temperature to describe social settings is also strong: 'warm and welcoming' service; being given the 'cold shoulder'; an establishment is 'cool and trendy'. This makes choosing appropriate thermal feedback to represent restaurants/bars difficult.

We wanted to understand how people would use temperature to convey their own experiences of bars and restaurants. This could relate to overall quality, or some other social or hedonic element. By looking at the relationships between thermal stimuli and meaning across different participants, we can see if there are any patterns that could be used to inform the design of thermal feedback. In this scenario, participants were asked to assign a temperature to places that they liked, disliked and were indifferent to, and provide an explanation of their temperature mapping. In this way, we would see how temperature was used across places of varying subjective quality.

Stimuli & Measures

This scenario was different from the others, as no predefined thermal stimuli were presented to participants. Instead, a list of 395 bars and restaurants from Glasgow were shown on the PC monitor. Participants were asked to search the list for six bars or restaurants that they had visited (all participants were from the city): two that they liked, two that were average and two that they did not like. Three data points were collected for each of the places: a user-selected temperature that represented his/her experience along with some text that explained their temperature choice. Finally, an indication of whether they thought the bar/restaurant was "good", "average" or "bad" was entered. The temperature was set through a slider, which changed the Peltier temperature in real-time. The possible temperatures were 22°C (far-left position) to 38°C (far right) at a resolution of 0.1°C. No numerical value for the temperature was shown.

Procedure

This experiment ran on a PC, with the participant sat at a desk, with a monitor, keyboard and mouse in front of them along with the Peltier devices and armrest. They were asked to search for each of the six bars/restaurants either by typing relevant text into a search bar or scrolling the alphabetical list. When an establishment was selected, the monitor showed a screen including the name and representative image along with the input slider, text box and drop-down menu. The participants were asked to set the temperature of the Peltiers in response to the question *"What temperature would you associate with this place?"* They were told they were free to base their feedback on any aspect of their experience. Once the temperature had been set, they were asked to type their reasoning for the stimulus, or what it represented, before indicating whether they thought the place was "good", "average" or "bad" from the drop-down menu.

Rating	★☆☆☆☆	★★☆☆☆	★★★☆☆	★★★★☆	★★★★★
Temp	25°C	27.5°C	30°C	32.5°C	35°C

Table 3: Recommended temperatures for conveying the star rating of bars and restaurants using thermal feedback

Results & Initial Discussion

The average (min and max) temperature attributed to the three overall quality ratings "good", "average" and "bad" were 34.8°C (24.5°C, 38.0°C), 29.8°C (25.6°C, 32.7°C) and 25.4°C (22.0°C, 35.2°C), respectively. A Friedman test on the average representative temperature showed a significant effect of quality (χ^2 (2)=38.89, $p<0.001$). *Post hoc* Bonferroni-adjusted ($p<0.0167$) Wilcoxon tests showed that all three ratings were significantly different from each other. There were no numerical values shown next to the slider in the GUI, yet participants naturally chose almost equally spaced values centred on the neutral temperature of 30°C (±5°C). The average temperatures attached to the good, average and bad establishments could be extended to represent the common 1-5 star ratings, as shown in Table 3. While ~35°C was chosen to represent "good", in some mobile scenarios, stimuli between 32-35°C might be difficult to tell apart [21,22], so the 4- and 5-star temperatures could instead be ~34°C and 38°C, to facilitate differentiation.

Analysing the attribution of temperatures to experiences, we identified five categories of meaning: 1) the overall rating or quality (e.g., warm for good; 76 responses), 2) hedonic/emotional meaning (e.g., nice atmosphere; 16 responses), 3) social aspects (e.g., friendly people; 10 responses) and representing the thermal properties of 4) the food/drinks consumed (e.g., hot coffee; 5 responses) or 5) the physical environment (e.g., cold at night; 8 responses). A summary of responses is in Table 4, including the temperatures participants chose to represent their experiences.

Response Category	**Subjective Experiences**	**Associated Temperature**
Quality or Rating	"Good restaurant, so warm temperature"	35°C
	"The food was ok so…neutral temperature"	30°C
	"Good food"	34.1°C
	"Not a fan. Not at all!"	26.3°C
Hedonic or Emotion	"Nice atmosphere"	36.5°C
	"[Poor] atmosphere"	24.6°C
	"[Good] because [the coffee] smells nice"	33.1°C
	"Really cute…place"	36.5°C
Social	"Friendly place…it's easy to meet people"	35.2°C
	"Hot girls" (interesting people)	34.1°C
Food, Drink	"Gave me a cold pie"	23.4°C
	"Very spicy…map…temperature [to] heat"	36.3°C
	"I mainly go for cold food"	24°C
	"They sell hot coffee"	33.1°C
Physical Environment	"I've only been at night when it's cold"	28.8°C
	"You get hot from dancing"	34.1°C
	"You'd go on a warm summer evening"	32.7°C
	"I went…in…winter and I was freezing"	22.2°C

Table 4: Selected subjective responses, grouped by category. Includes the thermal feedback representing the experience.

It was clear that the vast majority of participants attributed the thermal stimulus to the overall quality or rating of the establishment, in line with the average temperatures for the overall ratings. Thermal feedback in a restaurant experience scenario appears to be used to represent the overall quality. Given the response consistency, there may be a strong link between thermal feedback and enjoyment. The second most common meaning attributed to thermal stimuli was hedonic or emotional comments. Nine responses commented on using temperatures to convey the atmosphere or mood in the place. For all responses, colder temperatures were used for poor atmosphere and warmer temperatures for good.

Several of the experiences relating to atmosphere were joined by social comments on the quality or *"friendliness"* of the service received from staff. Generally, the quality of service correlated with the overall quality rating, so warm temperatures were used for good service and cold for bad. The quality, atmosphere and social attributes all gel with the results from the two previous scenarios, where warmth is characteristic of physical or social warmth. Eight responses used the temperature to represent a corresponding environmental temperature, while the remaining responses related to the physical characteristics of the food, such as poor cold food or nice hot coffee.

It is noteworthy that the two restaurant experience categories that represent real, concrete features (food or environmental responses) numbered so few, while the vast majority fall under abstract or subjective categories (quality, hedonic, social). This suggests that, at least in relation to a restaurant experience, thermal feedback is not seen a descriptive or representative signal, one that conveys an actual state. It is an excitatory, emotional signal for conveying high-level impressions. Thermal feedback has been shown to be effective and convincing in mimicking object/surface thermal properties in virtual reality [5,16]. However, this was not a common usage for thermal feedback by participants here.

DISCUSSION & FEEDBACK GUIDELINES

In all four scenarios there was a strong uniformity in the participants' interpretation of thermal feedback, in slight contrast to previous research [14,18]. In general, our results suggest that warm feedback represents 1) the presence of life or activity and 2) emotional positivity, while cold feedback represents the opposite: 1) the absence of people and activity and 2) emotional negativity. When asked how recently a person had been active on social media, warm temperatures (34-38°C) were consistently interpreted as meaning very recent activity (no more than a few hours), while cold temperatures (22-26°C) were interpreted as meaning much older activity (weeks and months). This could be interpreted as a *temporal* social closeness, rather than a *physical* social closeness. This relationship of warmth to presence and cold to absence was also found in the office availability scenario. There was strong agreement that warm temperatures (34-38°C) indicated presence in the office and cold temperatures (22-26°C) indicated absence. Stronger warmth was also interpreted as meaning the person had a higher level of unavailability ("Extremely busy", compared to "Available for short times"), while stronger cold meant that the individual was away for longer. Another view on the social media and presence data could be that a person's activity level could be indicated along an axis from *no activity* (cold 22°C) to *most activity* (warm 38°C).

The results from the first three scenarios throw up an interesting comparison to previous research on the perception or identification of predefined thermal changes for conveying information in HCI. Research has shown that our ability to identify a virtual material from its thermal conductivity (change in temperature over time) alone can be poor, ranging from 16% to 100%, even when choosing between only 4-5 possible materials [5,6,7]. The identification of more explicitly structured thermal feedback (e.g., 2 to 3 set levels of warmth) is more accurate, at around 75-100% for 4-5 levels or alternatives [19,21,22]. Both the social media and content deletion scenarios had reliable patterns of responses. Excluding the neutral temperatures (30-32°C), each of the remaining 7 stimuli in the social media scenario had a different pattern of time frames being attributed to them. The deletion scenario showed significant differences in perceived magnitude between all but the 3 lowest temperatures, which suggests participants could differentiate 7 levels. This includes 4 different levels of warmth, where identifying only 2 has been challenging previously [21,22].

We did not test absolute identification, and participants were not told how many stimuli were used (or how far apart they were). However, it is unlikely that participants could identify each stimulus uniquely, due to the lack of explicit structure and the relatively large number of stimuli. What our results suggest, however, is that participants were still able to reliably perceive, and appraise differently, up to 7 levels of thermal change, indicating an inherent or subconscious appreciation of several extents of temperature change. This is a larger number than users have been able to consciously identify in previous research. The implication for thermal interface design is that, while users may not be able to deliberately identify many levels, perhaps due to comparison-induced uncertainty, they can be relied on to appreciate and make use of several different thermal states. In this case, thermal feedback may be better suited as an ambient or supplemental feedback method than a primary means of conveying specific information.

The restaurant experience scenario resulted in a narrower range of responses than we had anticipated, given the open-ended nature of how participants could choose the stimuli and assign meaning to them. 66% of all the meanings attached to the thermal stimuli related simply to the overall quality of the establishment. This strong uniformity suggests that, within the restaurant scenario, thermal feedback has a clear meaning and widespread interpretation. This is helpful for interaction designers as they can be confident in how the feedback will be interpreted: the higher rated the place, the hotter the feedback should be. Warmth (>30°C) could also be used to indicate a friendly or social atmosphere. It should be noted that the framing of the scenario, asking participants to choose good, average and bad places, could have unintentionally guided them towards a quality-based perspective. However, they were explicitly told that the stimulus could represent any aspect of their experience.

Feedback Design Guidelines

High degree of common interpretation

Participants interpreted thermal feedback in a consistent manner, with very similar views on the meaning of both warm and cool changes. Previous research has generally shown a similar consistency [14,18]. Therefore, despite the variable nature of thermal perception [9,23], thermal feedback designs can be created with a good degree of reliability in how they will be interpreted by different users.

Warmth means social and physical presence

Warmth (>32°C) should be used to convey the physical or social presence of other people, while cool (<30°C) should be used to convey the absence of people or usage. Greater or lesser degrees, respectively, of presence, activity and positivity can be conveyed through increasing the extent of heat or cold. This could include the digital presence of someone online, the physical presence of a friend nearby (using location aware services) or a friendly, social occasion, such as calendar events. A number of different social/spatial states could also be conveyed through thermal feedback: spatial proximity in navigation, website traffic or the number of social media posts on a topic.

Temperature maps to quality

The quality or rating of content can be conveyed through temperature, with cool (~22-25°C) indicating the lowest quality and warmth (~35-38°C) indicating the highest.

Thermal feedback can convey risk-related status

When a user attempts to alter data, thermal feedback can reliably convey how used (and perhaps important) the content is. This could help to impose upon the user the risk of permanent alterations to their information.

Users can appreciate multiple levels of ambient feedback

While unique identification of thermal stimuli is challenging, users are able to appraise and make use of multiple (in our case up to 7) different feedback temperatures. Feedback designs can therefore reliably utilize different temperatures, but should do so in an ambient or supportive manner.

CONCLUSIONS

This paper presented studies that investigated how participants interpret thermal feedback in the context of four different types of information: online social activity, physical presence, digital content usage and experiential content. Participants reliably interpreted multiple levels of thermal feedback in consistent ways, attributing similar meanings to warm and cool changes. Warmth indicated 1) more recent social media activity, 2) physical presence and busyness, 3) higher content use and 4) positive experiences at bars or restaurants. Cooler stimuli conveyed the opposite. The results provide new insight into what inherent meaning thermal feedback has in HCI contexts and our design guidelines suggest how it might be effectively utilized in interfaces.

REFERENCES

1. Gallo, S., Cucu, L., Thevenaz, N., Sengul, A., and Bleuler, H. Design and control of a novel thermo-tactile multimodal display. *Proc. HAPTICS 2014*, pp 75–81.
2. Gooch, D. & Watts, L. Communicating social presence through thermal hugs. *Proc. SISSE 2010*, pp 11–19.
3. Halvey, M., Henderson, M., Brewster, S., Wilson, G., and Hughes, S. Augmenting Media with Thermal Stimulation. *Proc. HAID 2012*, pp 91–100.
4. Halvey, M., Wilson, G., Brewster, S., and Hughes, S. "Baby It's Cold Outside": The Influence of Ambient Temperature and Humidity on Thermal Feedback. *Proc. CHI 2012*, pp 715–724.
5. Ho, H.-N. and Jones, L. Contribution of thermal cues to material discrimination and localization. *Perception & Psychophysics 68*, 1, pp 118–128.
6. Ho, H.-N. and Jones, L. Development and evaluation of a thermal display for material identification and discrimination. *ACM TAP,* 4 (2) 2007, Article 13.
7. Ino, S., Shimizu, S., Odagawa, T., *et al*. A tactile display for presenting quality of materials by changing the temperature of skin surface. *Proc. IEEE Robot and Human Communication*, pp 220–224.
8. Iwasaki, K., Miyaki, T., and Rakimoto, J. AffectPhone: A Handset Device to Present User's Emotional State with Warmth/Coolness. *Proc. BIOSTEC 2010*, pp 1–6.
9. Jones, L. and Berris, M. The Psychophysics of Temperature Perception and Thermal-Interface Design. *Proc. HAPTICS 2002*, pp 137–142.
10. Kenshalo, D., Holmes, C., and Wood, P. Warm and Cool Thresholds as a Function of Temperature Change. *Perception & Psychophysics 3*, 2A, pp 81–84.
11. Kenshalo, D., Nafe, J., and Brooks, B. Variations in Thermal Sensitivity. *Science 134*, 1, pp 104–105.
12. Lautenbacher, S. and Rollman, G. Sex differences in responsiveness to painful and non-painful stimuli are dependent upon the stimulation method. *Pain 53*, 1, pp 255–264.
13. Lee, W. and Lim, Y. Thermo-Message: Exploring the Potential of Heat as a Modality of peripheral Expression. *Proc. CHI 2010 Ext. Abs.*, pp 4231–4236.
14. Lee, W. and Lim, Y. Explorative research on the heat as an expression medium: focused on interpersonal communication. *Personal and Ubiquitous Computing 16*, pp 1039–1049.
15. Nakashige, M., Kobayashi, M., Suzuki, Y., *Tamaki, H., and Higashino, S*. "Hiya-Atsu" media: augmenting digital media with temperature. *Proc. CHI 2009 Ext. Abs.*, pp 3181–3186.
16. Richter, H., Hausen, D., Osterwald, S., and Butz. Reproducing Materials of Virtual Elements on Touchscreen using Supplemental Thermal Feedback. *Proc. ICMI 2012*, pp 385–392.
17. Salminen, K., Surakka, V., Raisamo, J., *et al*. Emotional Responses to Thermal Stimuli. *Proc. ICMI 2011*, pp 193–196.
18. Suhonen, K., Muller, S., Rantala, J., *et al*. Haptically Augmented Remote Speech Communication: A Study of User Practices and Experiences. *Proc. NordiCHI 2012*, pp 361–369.
19. Wettach, R., Behrens, C., Danielsson, A., and Ness, T. A thermal information display for mobile applications. *Proc. MobileHCI 2007*, pp 182–185.
20. Williams, L. and Bargh, J. Experiencing physical warmth promotes interpersonal warmth. *Science 322*, 1, pp 606–607.
21. Wilson, G., Brewster, S., Halvey, M., and Hughes, S. Thermal Icons: Evaluating Structured Thermal Feedback for Mobile Interaction. *Proc. MobileHCI 2012*, pp 309–312.
22. Wilson, G., Brewster, S., Halvey, M., and Hughes, S. Thermal Feedback Identification in a Mobile Environment. *Proc. HAID 2013*, Article 2.
23. Wilson, G., Halvey, M., Brewster, S., and Hughes, S. Some Like it Hot? Thermal Feedback for Mobile Devices. *Proc. CHI 2011*, pp 2555–2564.