Synestouch

Haptic + Audio Affective Design for Wearable Devices

Pablo Paredes, Ryuka Ko, Arezu Aghaseyedjavadi, John Chuang, John Canny
[pablo.paredes], [ryukako], [arezua], [chuang], [canny] @berkeley.edu
University of California – Berkeley

Linda Bäbler
l.m.babler@student.tue.nl
Eindhoven University of Technology

***Abstract*— Little is known about the affective expressivity of multisensory stimuli in wearable devices. While the theory of emotion has referenced single stimulus and multisensory experiments, it does not go further to explain the potential effects of sensorial stimuli when utilized in combination. In this paper, we present an analysis of the combinations of two sensory modalities -- haptic (more specifically, vibrotactile) stimuli and auditory stimuli. We present the design of a wrist-worn wearable prototype and empirical data from a controlled experiment (N=40) and analyze emotional responses from a dimensional (arousal + valence) perspective. Differences are exposed between "matching" the emotions expressed through each modality, versus "mixing" auditory and haptic stimuli each expressing different emotions. We compare the effects of each condition to determine, for example, if the matching of two negative stimuli emotions will render a higher negative effect than the mixing of two mismatching emotions. The main research question that we study is: When haptic and auditory stimuli are combined, is there an interaction effect between the emotional type and the modality of the stimuli? We present quantitative and qualitative data to support our hypotheses, and complement it with a usability study to investigate the potential uses of the different modes. We conclude by discussing the implications for the design of affective interactions for wearable devices.**

Keywords*— *Haptics, Vibration, Vibrotactile, Multimodal, Multisensory, Sound, Auditory, Stimulus, Wearable, Affective

I. INTRODUCTION

As wearables become increasingly equipped with multiple sensors and actuators, it is expected that perceptual experiences will be influenced not only by how each of these sensors and actuators are used individually, but also by their various combinations. Individual audio signals have been studied extensively as carriers of emotional content [2], [22]. Hertenstein et al. [9] have also shed light on the emotional expressivity of haptic signals. However, the affective response of these sensory stimuli is understudied.

This paper examines such multisensory combinations. More specifically, we study the perception of emotions when triggered by sounds and/or vibrotactile stimuli. We choose a wrist-worn device mainly due to its large adoption as a wearable, despite the complexity of generating haptic stimulus due to its limited contact with the human body.

We evaluate emotion from a dimensional perspective, based on the continuous measurement of their emotion components: arousal and valence [9], [16]. In general, we expect that the combination of multiple stimuli should generate different responses depending on the intensity and the value of their arousal and valence components. Our main research question is: *When haptic and auditory stimuli are combined, is their combination linear in nature?*

A preliminary qualitative assessment helped illuminate the nature of emotionally non-matching haptic and audio stimuli combination. The combination of non-matching stimuli (i.e. pertaining to two different quadrants) generated responses of surprise and disgust, which indicate that results will most likely fall in the high arousal, low valence quadrant. With this information, we formulated the following hypothesis:

H. When combining sound and vibrotactile stimuli, the following interaction effects will be observed:

H.1 auditory stimuli will be dominant in both axes (valence and arousal).

H.2 haptic stimuli will modify (reduce or enhance) the effect of sound.

To test this hypothesis, we built a wearable prototype made of two eccentric rotating mass (ERM) vibration motors to induce two haptic phenomena known as "apparent tactile motion" and "phantom tactile sensation" [12], in order to generate some of the affective gestures described by Hertenstein et al. [9]. We performed a controlled lab experiment with 40 users (52% female and 48% male) balanced across conditions. We controlled the order in which the stimuli were presented to the user. Our results show that there is no interaction between stimuli modes and their emotional types (arousal / valence). Single-mode stimuli showed a higher valence for sound (vs. vibrotactile). Multimodal interactions did not reflect a clear dominance of sound, which was modulated by haptic stimuli. However, the lack of clear emotional expressivity of the single modes stimuli affected the combined modes. Furthermore, qualitative data supported these findings.

II. BACKGROUND

A. Circumplex Model of Affect

The Circumplex Model of Affect (CMA) by Russell [16] essentially represents emotions as a combination of two dimensions: arousal (ranging from sleep to high arousal) and valence (ranging from displeasing to pleasing). As depicted in Figure 1, emotions are organized in a two-dimensional space. The horizontal axis represents the feelings/valence (pleasure-displeasure), while the vertical axis represents the degree of arousal (excited-sleep). When looking at the CMA, it becomes apparent that each of the four different quadrants contains similar emotions. The upper right quadrant is themed around emotions of excitement or happiness, the upper left represents distress or anger, the lower left is representative of depression or sadness, and the lower right revolves around feelings of contentment or relaxation.

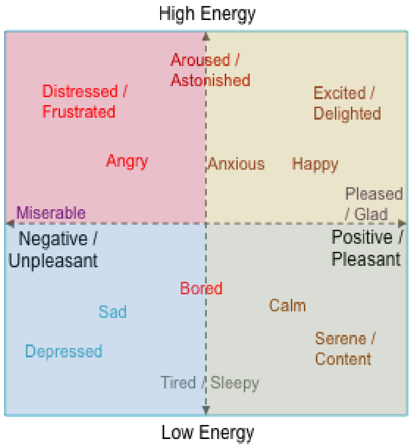

Fig. 1. Circumplex Model of Affect by Russell [16].

B. *Auditory Affective Expressivity*

Affective expressivity of sounds is well studied and documented in the International Affective Digitized Sounds (IADS-2) library developed by Bradley and Lang [2]. The IADS-2 is a set of standardized, emotionally-evocative, internationally accessible sound stimuli that consists of a total of 167 sounds, each of which is characterized by its mean values for arousal and valence.

C. *Haptic Affective Expressivity*

A comprehensive study of affective expressivity of human touch, performed by Hertenstein et al. [9] evaluates twelve emotions divided into three groups: (a) 6 emotions based on Ekman's traditional studies of emotion [5] (anger, fear, happiness, sadness, disgust, and surprise), (b) 3 prosocial emotions (love, gratitude, and sympathy) [17], expected to have higher expressivity via touch, and (c) three self-focused emotions (embarrassment, pride, and envy) [21], which play as counterparts to the prosocial ones. In addition, by observing how participants expressed emotion through touch, Hertenstein et al. [9] could determine the types of touch people used to communicate specific emotions. Table I shows a subset of the types of touch selected for this paper.

TABLE I
TOP FOUR HAPTIC PATTERNS USED FOR THE EXPRESSION OF EMOTIONS [9].

Emotions	**Types of touch – in order of relevance**
Anger	Hitting, Squeezing, Trembling.
Fear	Trembling, Squeezing, Shaking.
Happiness	Swinging, Shaking, Lifting.
Sadness	Stroking, Squeezing, Lifting.

D. *Multisensory Affective Response*

Haptics and visual stimuli interaction has been recently studied showing a dominance of the visual stimulus, despite the different modulations of the haptic signals performed by the authors [1]. No additional work in other types of multimodal affective responses have been reported to date.

III. PRIOR WORK

A. *Audio and Haptic Research*

Haptic and audio interaction workshops [8] conducted over the course of the past decade have studied the use of audio and haptic interaction in fields as diverse as music, interfaces, and communications [3], [18], [19], with a focus on the design and evaluation of novel multimodal interactions. However, the use of haptic technology to support affective computing is barely studied. Within the literature, we find that haptic stimulus could help reduce the amount of sensorial or cognitive load when implementing calming technologies [13], [14]. Another line of study is the design of frameworks to examine the use of haptics primitives such as distance, surface contact, time exposure, movement, and oscillations, among others [4], [7].

B. *Vibrotactile Interfaces*

Israr and Poupyrev [12] explored the use of two vibration sources to create the illusion of apparent tactile motion, which could be used to mimic the touch of a human stroke, or the phantom tactile sensation, which could be used to mimic a poke (see Figure 2).

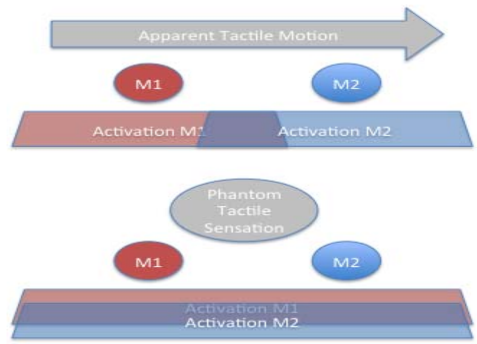

Fig. 2. Basic haptic effects with two vibrating motors as described in [12].

Recently, Huisman et al. [11] evaluated interactions based on a forearm sleeve used to express certain movements such as a poke, a hit, pressing, squeezing, rubbing, and stroking. Richter et al. [15] explored scenarios where the phantom tactile sensation can be recreated using different actuators with direct applications in ubiquitous computing scenarios.

IV. EXPERIMENT DESIGN

A. *Vibrotactile Wearable Hardware*

We built a wristband made of two Velcro strips. In between the strips, we fixed two small 5V DC ERM vibration motors with a variable speed ranging from 0 RPM to 9000 RPM. Figure 3a shows that the motor is directly connected to a pulse-width modulation (PWM) pin of the microcontroller. The motors were placed at a distance of 1.1 inches apart and they were oriented to be parallel to the bones in the forearm. Figure 3b shows the placement of the motors on a human forearm.

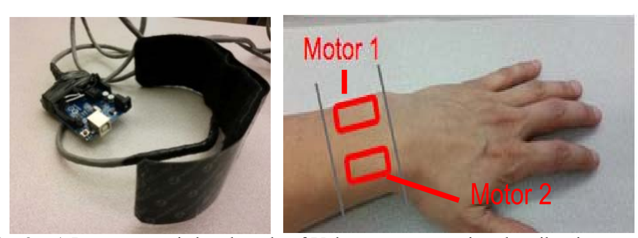

Fig. 3. a) Prototype wristband made of Velcro straps covering the vibrating motors and attached to Arduino Uno. b) Motor placement (1.1 inches apart).

With the location of these motors, we were able to recreate the "apparent tactile motion" effect [12] by using an overlap time of 200ms between the end of the stimulus in one motor and the beginning of the stimulus on the other motor. We were also able to recreate the "phantom tactile sensation" effect [12] by activating both of the motors in parallel.

B. Stimuli Generation

Table II shows a representation of the CMA with the selected vibrotactile and sound stimuli for each quadrant. We gave them short labels (happy, relaxed, sadness, anger) for easier identification purposes. Sounds were played on a speaker under the chair's armrest at a uniform intensity, which was calibrated to the user's comfort, while vibrotactile intensity depended on the type of interaction. The stimuli had a duration of 6.5 seconds, which seemed to be a good time to elicit enough emotional information based on our previous pilot observations.

We selected our touch patterns based on the touch primitives described by Hertenstein et al. [9]. We chose a "hitting" touch to represent the high arousal + low valence (anger) quadrant, a "swinging" touch for the high arousal + high valence (happiness) quadrant, and a "stroking" touch for the low arousal + low valence (sadness) quadrant. Hertenstein et al. [9] did not research emotions relating to the low arousal + high valence quadrant (relaxation), so we chose a low intensity long touch based on the idea of a "relaxing" massage.

We translated our selected touch patterns into vibrotactile stimuli based on the implementations described by Huisman et al. [11]. Touch patterns were created based on different combinations of the duration and the intensity of vibration. The "hitting" touch was represented by a poke using the "phantom tactile sensation" effect with a short duration and high intensity vibration. The "swinging" touch utilized the "apparent tactile motion" effect, by using quick left to right and right to left strokes of moderate intensities. The "stroking" touch also used the "apparent tactile motion" effect with a unidirectional stroke of moderate intensity. Lastly, the "massage" touch used a long duration and low intensity activation of both motors.

TABLE II
CMA WITH HAPTIC AND SOUND MAPPINGS PER QUADRANT.

Haptic: Hitting → Hit High Intensity (5V) Short Burst (100ms both motors) "Phantom Tactile Sensation" **Sound → Alarm (IADS)** Valence = 4.3, Arousal = 6.99 Short label: "Anger"	**Haptic: Swinging → Strokes** Moderate Intensity (4V) Left-Right and Right-Left fast strokes (190ms, 50ms overlap) "Apparent Tactile Motion" **Sound: People Laughing (IADS)** Valence = 7.78, Arousal = 5.942 Short label: "Happy"
Haptic: Stroking → Stroke Moderate Intensity (4V) Left-Right slow stroke (380ms, 100ms overlap) "Apparent Tactile Motion" **Sound: Violin (own selection)** Valence = N/A, Arousal = N/A Short label: "Sadness"	**Haptic: Massage* → Press** Low intensity (2.5V) Long vibration (1500ms) "Phantom Tactile Sensation **Sound: Harp (IADS)** Valence = 7.44, Arousal = 3.36 Short label: "Relaxed"

We selected the sounds from the IADS-2 database based on the average arousal and valence values. Unfortunately, IADS-2 does not have a good low valence + low arousal sound. So, we queried an open source sound database [6] for a "sadness sound" and settled with the sound "Sad Violin."

C. Conditions

We conducted an experiment focused at evaluating the interaction between single and multimode stimuli as well as the differences between haptic and auditory-only stimuli. As it can be seen in Table III, in essence we designed a 2x2 experiment in order to counterbalance the conditions to reduce ordering effects.

TABLE III
CMA WITH HAPTIC AND SOUND MAPPINGS, ONE MAPPING FOR EACH QUADRANT.

	Single Mode (first)	**Combined Mode (first)**
Haptic (first)	**A)** 2x 4-Haptic stimuli 2x 4-Sound stimuli 2x 16-combined stimuli	**B)** 2x 16-combined stimuli 2x 4-Haptic stimuli 2x 4-Sound stimuli
Auditory (first)	**C)** 2x 4-Sound stimuli 2x 4-Haptic stimuli 2x 16-combined stimuli	**D)** 2x 16-combined stimuli 2x 4-Sound stimuli 2x 4-Haptic stimuli

Participants were asked to assess arousal (low to high energy) and valence (unpleasant to pleasant feelings) at the beginning of the experiment and after each stimulus using Likert scales from 1 to 9. The single-mode blocks consisted of 4 different auditory or 4 different haptic stimuli, all randomized. The multimodal blocks consisted of 16 randomized stimuli made of the combination of the 4 vibrotactile and the 4 sound stimuli. After each block there were short 30 second breaks. At the end of the experiment, they were asked to provide open text qualitative feedback about their experience and complete a usability questionnaire.

We gathered two (repeated) runs for each condition in order to reduce novelty effects. Weierich et al. [20] describe the effects of novelty in the amygdala as being very similar to those triggered by stimuli with high (vs. low) arousal and negative (vs. positive) valence responses. A simple average of the values of the two (repeated) runs can reduce novelty effects by improving the estimated value of each metric.

V. RESULTS

A. Quantitative Analysis

Our quantitative analysis follows a parsimonious procedure. It starts with a three-way hierarchical (nested) ANOVA that compares **stimulus *mode*** (haptic vs. auditory), **emotional *type*** (Low/High valence vs. Low/High arousal) and **emotional *match*,** a nested variable that differentiates between matching stimulus types (same valence levels and same arousal levels) versus non-matching stimulus types (different valence levels and/or different arousal levels) for the combined conditions. We created two models, one for arousal and one for valence.

1) Pre-processing and Data Visualization: To compare our Likert scale measurements, we normalized the intercepts across subjects by subtracting the baseline value measured at the beginning of the study. The arousal and valence scales that ranged from 1 to 9 were converted to scales that ranged from -8 to 8 with a common zero value. Figures 4 and 5 show the mapping onto the CMA of the haptic and sound stimulus values respectively. As it can be observed, there is very little difference between the centroids and its distributions, and most of the values were below zero for both arousal and valence. Having most of the values below zero indicates that in general, people did not find the experience exciting (low energy/arousal) and that it generated some discomfort (unpleasant feelings).

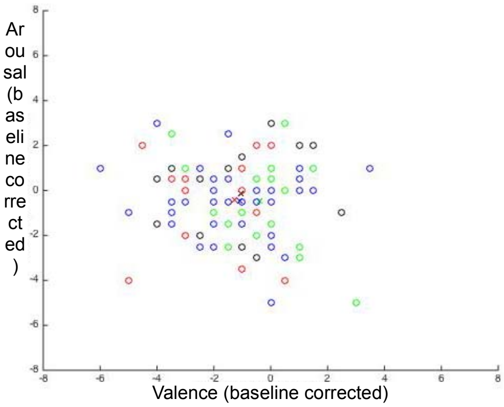

Fig. 4. Haptic stimulus affect map: (o) represents average points for all 40 users; (x) represents the centroids; colors represent the different stimuli: black = happy, red = anger, blue = sadness, green = relaxed.

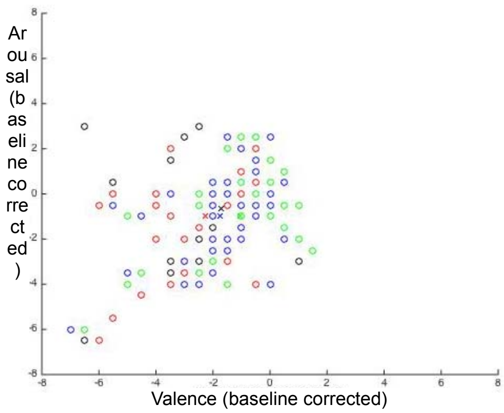

Fig. 5. Sound stimulus affect map: (o) represents average points for all 40 users; (x) represents the centroids; colors represent the different stimuli: black = happy, red = anger, blue = sadness, green = relaxed.

2) Arousal Data Analysis: Results from a three-way hierarchical (nested) ANOVA for mode, type and match(type) shown in Table IV reflected neither significant effects from any interaction nor main effects. It also showed that the nested variable match emotional type had no effect in the analysis.

TABLE IV
THREE-WAY ANOVA FOR AROUSAL.

Source	Sum Sq.	d. f.	Mean Sq.	F	Prob>F
Type	45.74	19	2.40724	0.81	0.698
Match(Type)	0	0	0	0	NaN
Mode	0.05	1	0.05	0.02	0.8969
Type*Mode	2.88	3	0.95833	0.32	0.8095
Match(Type)*Mode	0	0	0	0	NaN
Error	2785.5	936	2.97596		
Total	2834.16	959			

The null hypothesis could not be rejected for emotional type, F(19)=0.81, p=0.698, nor for stimulus mode, F(1)=0.02, p=0.8969 nor for the interaction emotional type * stimulus mode F(3)=0.32, p=0.8095. This indicates that arousal levels could not be differentiated across the different single-mode and combined-mode stimuli. The lack of expressivity along the arousal axis makes it impossible to establish a concrete interaction effect between the different stimuli. However, it is possible to say that the combined stimulus remained close to the original single-mode stimulus. Therefore, for arousal neither H.1 nor H.2 could be accepted.

3) Valence Data Analysis: Results from a three-way hierarchical (nested) ANOVA for mode, type and match(type) shown in Table V similarly reflected that there were no interaction effects between emotional type * stimulus mode F(3)=1.73, p=0.1593 and again, the match-type nested factor had no effect in the analysis. However, in this case both stimulus mode and emotional type had significant effects.

TABLE V
THREE-WAY ANOVA FOR VALENCE.

Source	Sum Sq.	d. f.	Mean Sq.	F	Prob>F
Type	166.42	19	8.7587	2.78	0.0001
Match(Type)	0	0	0	0	NaN
Mode	23.11	1	23.1125	7.33	0.0069
Type*Mode	16.36	3	5.4542	1.73	0.1593
Match(Type)*Mode	0	0	0	0	NaN
Error	2951.53	936	3.1533		
Total	317.42	959			

Figure 6 shows the stimulus mode's observed main effect, F(1)=7.33, p=0.0069. We performed a multi-comparisons test with Bonferroni's correction (p=0.0167). In this case, sound has a higher valence (M=-0.78, SE=0.1417) than haptic (M=-1.32, SE=0.1417) or combined (M=-1.34, SE=0.1417). This showcases a potential dominance of sound over other modes.

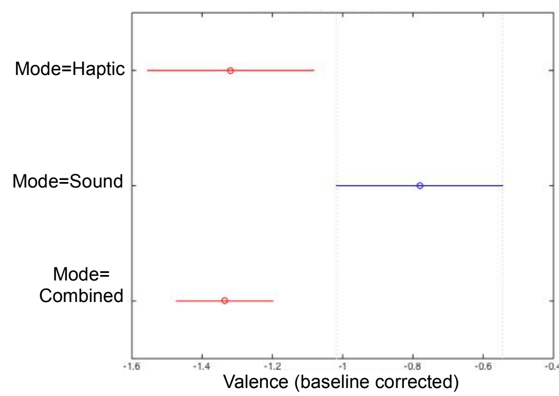

Fig. 6. Multiple comparisons (with Bonferroni's correction). Sound mode's valence is significantly different than Haptic mode or Combined (Haptic + Sound) mode.

With this information, we compared the emotional types for single-mode with the combined modes. We observed that the main effect was significant, F(19)=2.78, p=0.0001. After performing multiple comparisons with Bonferroni's correction (p=0.0025), it can be appreciated in Figure 7 that the only emotional type that was statistically different to the others was the combined Happy (haptic) + Anger (sound) which had lower valence than the Relaxed (sound only) stimulus and the matching Relaxed (haptic) + Relaxed (sound).

To further disambiguate which single-mode stimuli had actual differences among their individual emotional types, we performed a one-way ANOVA for vibrotactile F(3)=2.07, p=0.106, which rendered non-significant. In the case of sound we observed a significant difference F(3)=2.9, p=0.037. Figure 8 shows a multiple comparisons chart with Bonferroni correction (p=0.0125), where it can be seen that the sounds that have statistically significant differences t(78)=-2.88, p=0.0052 are Anger (low valence, high arousal) M=-2.24, SD=1.89 and Relaxed (high valence, low arousal) M=-1.06, SD=1.8.

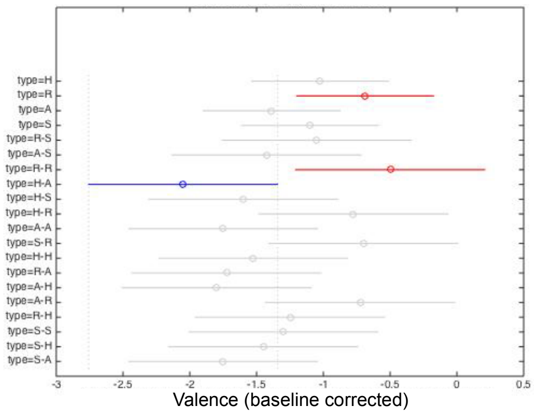

Fig. 7. Multiple comparisons (with Bonferroni's correction). Happy haptic (High Valence + High Arousal) combined with Anger sound (Low Valence + High Arousal) was significantly different.

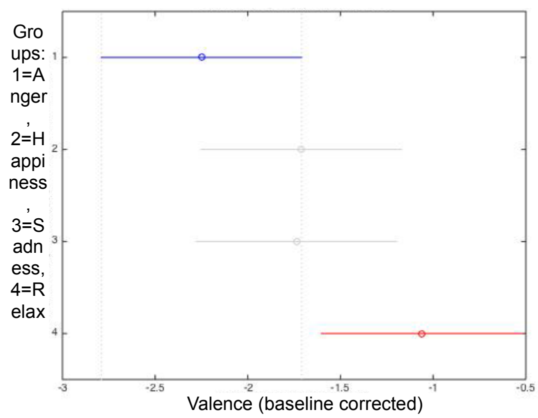

Fig. 8. Multiple comparisons (with Bonferroni correction p-0125). Only Anger and Relaxed present significantly different valences.

With this new information, it seems plausible that the reason Happy (haptic) + Anger (sound) is different than Relaxed (haptic) + Relaxed (sound) and Relaxed (single-mode) is due to the effect of sound. However, this is non-conclusive, and given the lack of statistical significant differences among the other types, we cannot assert that sound indeed had stronger dominance. We could, on the contrary infer that the power of the single-mode Anger and Relaxed sounds did not prevail across all the other combinations containing Anger or Relaxed sounds and that the haptic seems to have a modulating effect when combined with sound.

B. Qualitative Analysis

Due to the lack of strong evidence from the quantitative assessment, it is relevant to analyze the open text data generated during the study.

1) Emotional Expressivity: Overall, participants noted sound as being more effective in conveying emotions, while haptic stimuli were harder to interpret.

"With haptics alone (without sound), it is much harder to determine what the haptics are trying to convey. It is very open to interpretation because there are many possible meanings of what the haptics mean. Sound alone (without haptics) can convey more meaning because specific sounds are easier to associate with things we experience/hear in everyday life."

"I thought it was more difficult to relate an emotion to haptics. For some of the combinations, the emotion I was feeling was not listed."

"Haptics is good for alerts (why I usually associate it with stress), but sounds are better for conveying emotions."

Haptic stimuli were largely associated with alarms, stress, and high levels of arousal:

"Haptics seemed inconsiderate, obnoxious and annoying most of the time. They feel appropriate for alerting someone to pay attention, but they didn't really convey any emotions besides stress when they were alone."

"When the vibrations were on, they were distracting, and depending on the amount and speed of the vibrations my 'antsyness' went up. I became uneasy."

"The haptics tended to make me feel more energized, but I don't think they affected my emotions as much."

Furthermore, the length and intensity of the haptic stimulus influenced whether it was perceived as pleasant or not. Most participants enjoyed long and soft or slowly pulsing stimuli while sharp, abrupt, or strong stimuli were perceived as aggressive or negative:

"The sharp haptics was almost annoying. Like someone poking me. The more constant rolling vibration was a lot more calming."

"I liked a more sympathetic longer vibration. Like it was saying 'I am really sorry that I am waking you up, but you asked for it'."

2) Perceptual Dominance: With regard to dominance, slightly more participants described sound as the dominant stimulus. When asked about stimulus dominance, 15 participants (37.5 %) chose "sound. " The remaining participants were equally spread between "haptics" or "both," with 12 participants (30.0 %) each. While sound dominated in conveying the emotional tone, participants mentioned haptics as producing a greater emotional nuance.

"However, the combination of sound and haptics makes the most impact. The combination of haptics with sound can give different interpretations of what the sound is meant to convey."

"The sound was the dominant way I determined the emotion, it overpowered the haptics. The haptics were what I used to determine which category out of the few that the sound could be related to. So for a sound that I felt could be fear, stress, or anger, the haptics would dictate which of those I placed it in."

"A soft haptic vibration soothed the sound of hard laughter (which could sound very annoying/intrusive). A pleasant sound coupled with an urgent, throbbing vibration made me angry (took away from what I wanted to feel). An urgent sounds with urgent/fast vibration made me feel like I wanted to flee rather than figure out what was wrong."

3) Combined Effect: The combination of stimuli was described to be enjoyable only when each haptic and sound stimuli were perceived to be matching in emotion. When the stimuli did not match, the combined trials were described as stressful.

"..., I would be inclined to say that the dominant frequency of the sound and the dominant frequency of the haptics need to be aligned. The misaligned sounds-haptics gave me the impression of bad design."

"There were combinations of sounds and haptics that angered or surprised me and I think it was due to a mismatch in the pairing. The laughter sound, when paired with the right vibration evoked happiness but when paired with another vibration made me feel angry or stressed."

"I expressed surprise at the early combinations of soothing/sympathetic music with vibration alerts, as they didn't seem properly paired... I didn't feel there was any emotional choice to match an alert or alarm..."

C. Usability Assessment

We investigated four usability metrics: *Potential Frequency of Use* (Daily, Weekly, Monthly and Other), *Likelihood of Usage* (1: least likely to 5: most likely), *Preferred Modality* (top-down best to worst rank), and *Perceived Dominance* (top-down best to worst rank).

A descriptive summary of the data is presented in Figure 8. The radius axes were all normalized (from 0 to 1) to create a graph that helps visualize the highest values or rankings outwards from the center of the radius plot.

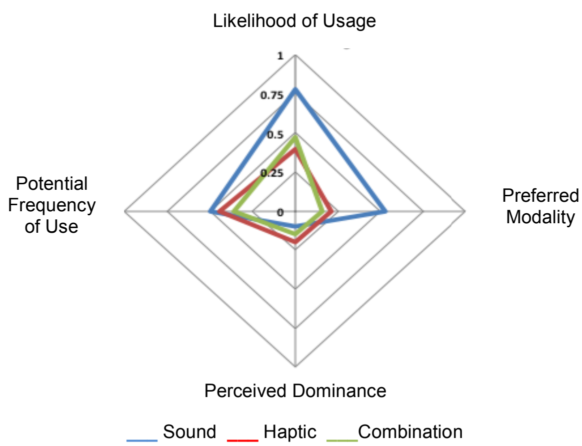

Fig. 8. Experimental Design for conditions 1 and 2.

From Figure 8 it can be observed that 51% (21 out of 40) participants preferred sound. Likelihood of use for sound (0.82) was also higher than haptic (0.42) and combined stimuli (0.58). Potential frequency of use and perceived dominance were essentially similar for all modes.

1) Gender Bias: Previous research showcases differences associated with gender and the perception of emotions [10]. We explored if there is a gender bias on the main usability metrics. 21 females and 19 males participated in the study. We found that only sound had a gender bias with respect to its likelihood of use (chi-squared = 44.8509, df = 12, $p<0.001$) and preference (chi-squared = 22, df= 2, $p<0.001$).

Going back to the qualitative data, both genders expressed that emotion is more easily expressed through sounds, while haptics tended to be felt as more annoying. Yet, in choosing which stimuli was more dominant in a multimodal condition, 8/19 (42%) males and only 4/21 (19%) females chose haptics as being dominant; as opposed to 4/19 (21%) males and 12/21 (57%) females that chose sound. Males cited the aggressive and alarming nature of haptics, and thus seemed to find haptics to be dominant over sounds even despite conceding to the greater emotional expressivity of sounds.

"I found it difficult to have a non alarming haptic - perhaps the lowest vibration level, longest pulse was closest (and least like any mobile phone alert-style vibration). Sound - even with the 3 or 4 limited clips used - allowed a richer range of emotions to be expressed." [NOTE: Subject chose "haptics" when answering which stimulus was more dominant when combined.]

Females similarly noted the alerting nature of haptics, but would still find sounds to dominate over haptic stimuli.

"Haptics is good for alerts (why I usually associate it with stress), but sounds are better for conveying emotions." [NOTE: Subject chose "sound" when answering which stimulus was more dominant when combined.]

2) Multimodal Communication: Table VI shows the preferred modes for a communication application, which indicates that users overall prefer to use messages with sound only or with the combined multimodal method.

TABLE VI
PREFERRED COMMUNICATION MODALITIES.

	Sound	Haptic	Combined	None
Counts	28	7	29	4
%	47,46	11,86	49,15	6,78

VI. Implications for Design

By looking at the quantitative results in isolation, one could conclude that there is little to be learned from the interaction between haptic and sound, mainly due to the poor resolution of the single-mode stimuli. However, we argue that the results of this paper should be explained from two perspectives: a) the lack of properly recognizable stimuli and b) the qualitative and usability data, which paints a more concrete picture of the situation. We concentrate these observation by noting a few implications for design for the design of multimodal (haptic and auditory) wrist-worn wearable devices:

A. Sound Stimuli Selection

Despite choosing sounds based on their previously documented affective content, our participants did not report substantial differences in arousal, and minor differences in valence. This, however, contrasts with the qualitative remarks, where people highlighted the higher expressivity of sound.

An appropriate design should consider personal preferences for sounds that elicit differentiable arousal and valence ranges. It can be argued as well that a dimensional view of affect is not an appropriate way to select different stimuli or to study its effects. A complement with a discrete view of emotions (Ekman's universal emotions [5]) could be of help.

B. Haptic Stimuli Selection

In the case of haptic stimuli, it was also evident that we did not achieve any affective differentiation. We based our design by mimicking anthropomorphic touch based on certain vibrotactile primitives such as poking, stroking, caressing, etc., which did not render proper affective responses. Perhaps a preliminary design challenge is to create a proper ontology of haptic stimuli with affective labels, where haptic primitives such as activation, decay, frequency, intensity, and duration, are used to describe affective primitives such as feelings, sensations, emotions, and moods.

C. Haptic as a Modulator of Sound

Overall, our findings show that despite a prevalence of sound as a better communicator of emotion, haptic stimuli do play an important role in the overall affective experience. When combined with sound, a properly matching haptic stimulus can serve as a booster for the emotion expressed by sound. However, choosing a haptic stimulus that does not match the emotional content of the sound could reduce or even eliminate the emotional expression of the sound, by turning it into disgust or stress.

D. Matching Primitives

Beyond affective coherence, the design of multimodal stimuli should also consider the effects of matching certain stimulus primitives such as frequency, intensity, rhythm, etc. Properly matching these primitives may not increase the overall affective experience, however, any mismatch could generate an adverse emotional reaction.

E. Gender Matters

The observed gender bias showcases the importance to tailor sounds, haptic, and multimodal affective expressions to different genders, especially if the interest is specifically to evoke or communicate emotions. For males, in the case of multimodal stimuli (sound + haptics), haptic could play a stronger emotional modulation role.

VII. Future Work

Our findings show the important relationships between sensory modes for affective perception. It is clear that more work needs to be performed to truly understand how devices can capture, but also express and help communicate emotions. One avenue is to focus on generating an ontology and a corresponding database of different haptic stimuli and the emotional responses that could potentially be generated by wearables and various Internet of Things devices. A complement to this analysis will be to understand other parts of the body where haptic stimuli could be applied. Finally, it would be relevant to investigate interactive scenarios where multimodal emotional expressivity can be of benefit, such as with autonomous vehicles, where mood and stress management will play an important role and where larger surfaces, such as the seat, could be used to improve the emotional expressivity of a haptic stimulus.

VIII. Conclusion

In the present study we have observed the potential affective expressivity of multimodal interaction of mixing haptic and auditory stimuli. We observed the lack of expressivity generated by the single-mode stimuli, despite existing theoretical foundations. We observed no statistically significant interactions, which contrasted with personal statements obtained through a qualitative assessment. Despite the lack of clarity from the quantitative analysis, it is observable that the interaction of the modes is not null, nor is dominated by one single-mode. Auditory stimuli tend to maintain certain dominance, but it is clearly not immune to the effects generated by the haptic stimuli. Finally, some design intuition is developed, which principally focused on contextual awareness, personalization, careful selection of single-mode stimulus, gender differences, and the potential need for mitigation of negative effects carried by mismatching stimuli.

Acknowledgments

We thank Prof. Kimiko Ryokai at the School of Information at UC Berkeley for her advice on selecting the research topic.

References

[1] Akshita *et al.*, "Towards Multimodal Affective Feedback: Interaction between Visual and Haptic Modalities," *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, pp. 2043-2052, 2015.

[2] M. Bradley and P. Lang, "The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective Ratings of Sounds and Instruction Manual," Technical Report B-3, Univ. of Florida, 2007.

[3] L. Brown and J. Williamson, "Shake2Talk: Multimodal Messaging for Interpersonal Communication," *Haptic and Audio Interaction Design*, vol. 4813, pp. 44-55, 2007.

[4] E. Cooper *et al.* "Building a Framework for Communication of Emotional State through Interaction with Haptic Devices," *Haptic and Audio Interaction Design*, vol. 6306, pp. 189-196, 2010.

[5] P. Ekman, "Facial expression and emotion," *American Psychologist*, vol. 48, no. 4, pp. 384-392, 1993.

[6] Freesound.org, "Freesound.org," *freesound.org*. Retrieved July 21, 2015 from: https://www.freesound.org.

[7] Y. Gaffary *et al.*, "Comparison of Statistical Methods for the Analysis of Affective Haptic Expressions," *Haptic and Audio Interaction Design*, vol. 7989, pp. 69-78, 2013.

[8] Proceedings of the HAID - Haptic and Audio Interaction Design International Workshops, 2006 - 2013

[9] M. Hertenstein *et al.*,"Touch communicates distinct emotions," *Emotion*, vol. 6, no. 3, pp. 528-533, 2006.

[10] A. Hofer *et al.*, "Gender differences in regional cerebral activity during the perception of emotion: A functional MRI study," *Journal of Neuroimage*, vol. 32, no. 2, pp. 854-862, 2006.

[11] G. Huisman *et al.*, "The TaSSt: Tactile sleeve for social touch," *2013 World Haptics Conference (WHC)*, pp. 211-216, 2013.

[12] A. Israr and I. Poupyrev, "Tactile brush: drawing on the skin with a tactile grid display," *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - SIGCHI '11*, pp. 2019-2028, 2011.

[13] P. Paredes and M. Chan, "CalmMeNow: exploratory research and design of stress mitigating mobile interventions," *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '11*, pp. 1699-1704, 2011.

[14] A. Pirhonen and K. Tuuri, "Calm Down – Exploiting Sensorimotor Entrainment in Breathing Regulation Application," *Haptic and Audio Interaction Design*, vol. 6851, pp. 61-70, 2011.

[15] H. Richter *et al.*, "Tactile feedback without a big fuss: simple actuators for high-resolution phantom sensations," *Proceedings of the 13th International Conference on Ubiquitous Computing - UbiComp '11*, pp. 85-88, 2011.

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

[17] S. Stürmer *et al.*, "Prosocial Emotions and Helping: The Moderating Role of Group Membership," *Journal of Personality and Social Psychology*, vol. 88, no. 3, pp. 532-546, 2005.

[18] V. Tikka and P. Laitinen, "Designing Haptic Feedback for Touch Display: Experimental Study of Perceived Intensity and Integration of Haptic and Audio," *Haptic and Audio Interaction Design*, vol. 4129, pp. 36-44, 2006.

[19] P. Vickers, "Ways of Listening and Modes of Being: Electroacoustic Auditory Display," *Journal of Sonic Studies*, vol. 2, no. 1, 2012.

[20] M. Weierich *et al.*, "Novelty as a dimension in the affective brain," *Journal of Neuroimage*, vol. 49, no. 3, pp. 2871-2878, 2010.

[21] J. Wood *et al.*, "Self-focused attention, coping responses, and distressed mood in everyday life," *Journal of Personality and Social Psychology*, vol. 58, no. 6, pp. 1027-1036, 1990.

[22] M. Zentner *et al.*, "Emotions evoked by the sound of music: characterization, classification, and measurement," *Emotion*, vol. 8, no. 4, pp. 494-521, 2008.