Exploring the Haptic Crayola Effect

Matthew Brehmer Jeff Hendy Inwook Hwang Andreas Sotirakopoulos

University of British Columbia

brehmer@cs.ubc.ca, jchendy@cs.ubc.ca, inux@postech.ac.kr, andreass@ece.ubc.ca

ABSTRACT

Haptic icons are short sequences of vibrations designed to provide meaningful information and feedback to the user. In order for the icons to be useful, strategies to help users to identify, distinguish, and recall them need to be developed. In this paper, we investigate the effect that naming haptic icons has on the accuracy with which users are able to identify, distinguish, and recall the icons. We have conducted a between-subjects experiment using 60 participants equally divided in three naming conditions: no names, pre-selected non-descriptive names, and self-selected names. The experiment examined the impact that different naming strategies had on the ability of participants to accurately match the icons to given stimuli as well as their ability to remember the names. Our results suggest that while many participants felt that the names were useful, naming did not actually have any significant impact on the accuracy with which users were able to match icons to targets. Additionally, participants reported that allowing them to choose the names for the icons enabled them to better remember and distinguish the icons than participants who were presented with icons given non-descriptive names. Despite this finding, the groups did not differ significantly in their ability to remember the icons.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Haptic I/O

General Terms

Human Factors.

Keywords

Haptics, Haptic icons, Naming haptic icons, user study, haptic perception

1. INTRODUCTION

The modality of touch presents a promising channel for conveying information without adding excessive cognitive load in environments where additional demand for cognitive resources cannot or should not be placed on vision or audition.

While a large proportion of haptic technology research is devoted to direct rendering of 3D environments, our focus is on the design and evaluation of haptic icons. Haptic icons are short sequences of vibrations designed to provide meaningful information and feedback on event notification, system and application state, or other contextual cues to the user. Following Weiser's principles for designing calm technology [31], haptic icons can serve to peripheralize information that would otherwise be distracting in visual or auditory modalities. This is reinforced by evidence demonstrating high recognition rates for peripheral haptic icons under varying amounts of workload [4]. Haptic icons and tactile feedback have also been shown to be beneficial to users in mobile contexts, such as by providing assistance during text entry tasks [2], resulting in subjective reports from users of a reduction in workload, and less need for additional graphical or auditory feedback. However in order for haptic icons to be effective, users need a way to clearly and efficiently map each icon to its meaning, and to retain this association in memory. Thus, it is imperative to develop strategies to help users differentiate and remember haptic icons.

We suspected that giving descriptive names to haptic icons would help users identify, differentiate, and remember the icons. This suspicion was inspired by the distinct and memorable names the Crayola Company has given to hundreds of colours of crayons, and crayon users' apparent ability to differentiate and remember the colours. Thus, we call our hypothesized effect of naming haptic icons on icon learnability the Haptic Crayola Effect.

The aim of this paper is to explore the Haptic Crayola Effect. In our experiment, 60 participants were equally divided into three naming conditions. In the first condition, no icon names were used. In the second condition, pre-selected non-descriptive icons names were provided by the experimenters. In the third condition, participants were able to freely choose names for the icons. Participants first completed a series of exercises to gain familiarity with the icons and, if applicable, their names. Then they completed a series of tasks in which they were asked to replicate a series of three target icons. The experiment concluded with a questionnaire to elicit subjective feedback on the conditions.

We hypothesized that participants in the self-selected names group would be more accurate in matching the icons to the targets than participants in the two other groups. However, our hypothesis was not validated by the data collected during the experiment as there were no significant differences between participants in the three groups in terms of accuracy or the time taken to match the stimuli to the target icons. Furthermore, we hypothesized that participants in the self-selected names condition would have increased ability to remember icon names than participants in the non-descriptive names group; participants in the self-selected names group did report greater ability to remember the names on the post-experiment survey. Finally we hypothesized that participants who claimed to remember the icon names after the end of the testing phase would perform better than those who reported that they could not remember the names; this hypothesis was not supported by the performance data collected during the experiment.

We suspect that the lack of significant differences in terms of performance between the three groups may have been due to inadequate exposure to the icon names in the initial part of our experiment, which was necessary in order to maintain consistency among the three conditions. It is also possible that haptic icon stimuli are susceptible to verbal overshadowing [22], a memory phenomenon related to the inadequacy of verbal thought to remember complex non-verbal stimuli. Our discussion section elaborates on these issues and presents promising directions for future research in haptic icon differentiability and learnability.

In the following section, we discuss literature relevant to learning and differentiating haptic stimuli as well as effects of naming in other domains. We then describe our experiment and its results. We finish with a discussion of the experimental results and possible alternative ways to help users differentiate and remember haptic icons.

2. RELATED WORK

Early research on the design of haptic icons was guided by principles of graphical and auditory iconic communication in computer interfaces. The association between a stimulus and meaning must be well understood, easily learned, and clearly discernible from the associations of other icons. For a summary of techniques for achieving individual icon saliency and learnability, regardless of modality, see [1]. It is also important to define representational and abstract icons; the former being a caricature of the real-world stimulus, while the latter requires the user to form a semantic link between the icon's meaning and an abstract stimulus previously devoid of real-world meaning. Types of icons are not necessarily a dichotomy, but rather fall on a continuous scale of abstraction.

2.1 Auditory icons

The use of auditory icons offloads demand on the user's visual attentional system. However without an obvious link between a sequence of beeps and its meaning, such associations must be learned. Gaver suggested that the more a stimulus's form depends upon its meaning, the easier it should be to learn [11]. Gaver proceeded to explore purely-representational atonal auditory icons. While such sounds can be familiar to many users, and are salient in isolation, the difficulty these icons face in practice occurs is in the context of large icon sets, when individual icon salience and discriminability are diminished. Blattner et. al. then proposed the use of tonal, abstract auditory icons with no intrinsic meaning [1], making use of hierarchical and transformational structural organization patterns. Tonal auditory icons would allow pitches and intervals to be discerned and generalized, properties absent from Gaver's atonal representational auditory icons. This technique allowed for ease of constructing families of auditory icons with related stimulus-meaning associations. The use of such short patterns, typically 2-4 pitches in a rhythmic sequence, would allow users to chunk larger icons into organizational units, supporting ease of learning. The great advantage of this hierarchical-transformational approach as a learnability heuristic was the ability to accommodate both novice and advanced users with simple and compound icons using short tonal phonemes with shared meaning associations. However, while being less susceptible to problems of scalability, this approach lacks a formal grammar for generating icons, and in practice tends to use only a small amount of the potential perceptual design space.

2.2 Haptic icons

Various psychophysical experiments have examined the haptic perceptual capabilities for finely distinguishing tactile and force feedback [14, 25], as well as the capability of such parameters to convey amounts of information [26]. The results of these experiments have motivated the design of haptic icons. As in the auditory domain, haptic icons may be metaphorical or abstract, again with implications for scalability and learnability. Metaphorical approaches have been shown to be effective for small stimulus sets [5, 4], in which icons are designed based on underlying application-related metaphors. Despite that in such cases the stimulus-meaning association may be intuitive and learned by users in a matter of minutes, the metaphor-driven strategy for creating the haptic icon set does not incorporate a systematic basis for determining relative stimulus salience or differentiability. As such, this strategy relies heavily on designer creativity in generating good metaphors, which can become particularly difficult for abstract concepts and applications with larger icon sets. The strategy further depends on whether sets of icons will be distinguishable and salient in combination [17]. Also of concern are unexpected pre-existing associations the user may have with a particular metaphor, making it difficult for users to learn a new meaning association to the stimulus [27].

For perceptually optimized abstract stimulus sets, there is the question of which vibrotactile or force feedback parameters to vary between individual icons, such as frequency, amplitude, waveform, and duration. Brown and Brewster aimed to determine the parameters that were both easily distinguishable and memorable, such that an icon which varies along these parameters may be easily labeled and remembered by a user [3]. They proposed that perceived rhythm and roughness were the most appropriate parameters to be used in the design of such icons, and declared a need to label these parameters to maximize icon recognition and learnability.

2.2.1 Haptic icon differentiability

Recent work has refined the process of creating abstract haptic icon sets with maximum differentiability through perceptual optimization [27]. For a review of how perceptually optimised icon sets are generated, see [16, 17]. This process has been used to perceptually validate haptic stimulus sets prior to assigning meaning to individual stimuli, for those which perceivable parameters vary on frequency, force amplitude, wave shape [17], rhythm [27], and melodic variants [23, 24, 29].

2.2.2 Haptic icon learnability

Enriquez demonstrated the impressive ability of users to learn deliberately arbitrary stimulus-meaning associations for frequency and waveform parameters in a small set of haptic icons felt through a haptic knob interface [7]. Participants recalled 73% of the correct associations at a later test, indicating the potential for increased perceivable information density of distinct haptic signals.

People are generally better at learning and retaining haptic stimulus-meaning associations with larger, perceptually-validated abstract haptic sets than they would believe or report. This was shown in an experiment in which participants learned a set of perceptually validated abstract haptic stimuli over a two-week period [8]. No significant difference in recall performance was noted between a group of arbitrarily-assigned stimulus-meaning pairs and a group of participant-defined stimulus-meaning pairs. Participants did, however, indicate a subjective preference for the condition in which they were permitted to make their own stimulus-meaning associations. It was noted that participants generally lacked confidence in their ability to recall stimulus-meaning pairs after a 2-week period, despite their demonstrated ability to do so. This gives support to the notion that it is not important for the designer to perceive a semantic association between stimulus and meaning, but rather for the learner to, as they will likely develop their own mnemonics for distinguishing and learning perceptually optimized stimuli.

Swerdfeger performed two longitudinal deployment studies using perceptually distinct icons [24]. A set of icons used for notification interrupted a foreground puzzle game played on a mobile handheld device equipped with the capability to produce haptic signals differing on rhythmic and melodic parameters. The first study utilized icons whose primary perceptual distinction varied on rhythm. The stimulus-meaning association of an icon did not have a significant effect on its learnability; it was observed that the assignment of similar meanings to perceptually distinct stimuli does not seem to have an obvious negative effect on learning. This is contrary to the theory in which perceptually similar stimuli should have similar meanings [1]. Participants generally reported that their efforts were concentrated on distinguishing similar stimuli, that making the stimulus-meaning association was fairly unimportant for learning the icons, and that the meaning essentially acted as a name for each icon. Many participants developed unique ad-hoc mnemonics for distinguishing between and learning the stimuli. In the second study, it was believed that a melodic component may have elicited a range of responses reflective of the user's relationship with melodic music. However, the evocation of a possible emotional response to a stimulus could have hindered the ability to form new abstract semantic associations. As such, many users struggled to perceive and learn melodic variation in the stimuli. It was concluded that purely rhythmic icons are currently more predictable in terms of learnability.

2.3 Stimulus naming and learnability

When haptic stimuli are designed perceptually, nuances differentiating physical stimuli may not be easily learned without a linguistic tag. In the aforementioned studies of haptic icon learnability, we hypothesize that participants may have used, albeit internally, descriptive names for the physical stimuli in order to remember the icons, as opposed to using the abstract associated meanings.

Brown and Brewster suggested a need for labeling the perceptual parameters of rhythm as a means of learning rhythmic haptic icons [3], independent of the meaning association. Gabrielsson carried out a psycho-acoustical experiment in which musician and non-musician participants were presented with a variety of simple rhythms from a variety of musical styles and asked to rate the stimuli along 92 adjective dimensions [10]. The results suggested that people primarily use three classes of words to describe and label rhythms. These classes are (1) structurally descriptive terms for the rhythms, describing note accentuation or pattern complexity; (2) terms that are characteristic of perceptual or movement actions, anthropomorphizing a rhythm stimulus, such as *floating*, *stuttering*, or *graceful*; and (3) emotional terms, such as *vital*, *dull*, *excited*, or *calm*. Musical-linguistic tags for timbre have also been precisely specified among musicians and acousticians: HLF von Helmholtz's timbre descriptions based on the presence of specific harmonics in the audio signal, originally published in 1877, are still used today [12]. Among musicians, the fine and consistent descriptions of rhythm and timbre are imperative for communication and training. This paper explores whether such terms are of similar benefit to non-musicians.

Descriptive words are used for taste stimuli as well, however a consistent language of terms beyond combinations of the taste primaries *salty*, *sweet*, *bitter*, *sour*, and *umami* does not exist other than idiosyncratic and anecdotal descriptions [13]. Without a larger standardized terminology, common concepts are seldom conveyed using words alone. An exception is among taste panelists, whose training includes a greater commonality in verbal taste concepts.

While a consistent terminology does not exist for odours among non-experts, odour memory and odour naming were shown to be strongly related, with consistency of odour label use at studying and testing to be the main predictor of odour memory among adults [15]. While personally generated labels for odours were somewhat arbitrary, their use demonstrated that odour recognition may evoke both perceptual and semantic memories. Contrary to these results, experiments comparing odour recognition ability between wine experts and novices revealed superior olfactory recognition by wine experts, but failed to show better performance of experts compared to novices on related verbal memory tasks, indicating that superior odour recognition at studying and testing in wine experts was not reliant upon linguistic and semantic memory for wine odour terms [20]. An explanation for these results is that trained wine experts' talents include the ability to rely solely on their sensory-perceptual memory abilities, ignoring inhibitory linguistic-semantic memory. This phenomenon is an example of verbal overshadowing [22], a memory illusion assumed to occur when individuals are required to name complex non-verbal stimuli, those which are difficult to capture in words, including wine odours [19] or faces [22]; in essence, one's perceptual memory is overshadowed by one's attempts to verbally communicate that memory following perceptual encoding. Novices are particularly susceptible to this effect, in which a potentially high level of perceptual expertise is inhibited by a modest level of verbal expertise.

If naming a rhythmic stimulus can promote long-term stimulus-meaning learning, and if such stimuli are not susceptible to verbal overshadowing, it becomes unclear which type of name to use; rhythms seem to be linguistically overloaded with structural, perceptual-anthropomorphic, and emotional words. We anticipate that naming and remembering haptic or auditory stimuli based on their physical characteristics may be analogous to naming and remembering colours. The process by which children learn colour terms involves several mental mappings [21]; the non-obvious links between words and stimuli are learned if they are used consistently, and eventually the abstraction of a perceived property of a stimulus to its colour word can be made. Unfortunately, there are very few widely-agreed upon colour names beyond the basic colour terms, and our ability for learning additional colours and their names is poor [30]. In addition, colours may change due to contrast and adaptation in the viewing environment, which further constrain our ability to learn precise colour differences. In controlled conditions, a trained colourist can distinguish between a million colours using a pairwise comparison task [28], and can remember as many as 1000 colour names, however many such names are idiosyncratic or imprecisely defined [30]. While many colour name choices are indeed idiosyncratic, the popular crayon manufacturer Crayola uses a wide array of names, many of which evoke a memory of a colourful physical material in the world that is familiar to the user, aiding in the process of discriminating and learning the set of crayon colours. The name of a stimulus is provided by the manufacturer; many of which adopted from the U.S. Commerce Dept.'s National Bureau of Standards book *Color: Universal Language and Dictionary of Names* [6].

How effective in comparison are individual users at naming the colours themselves, and later remembering these name-stimulus associations? This question can also be asked of haptic stimuli. To our knowledge, no previous experimental research has used explicit names for individual haptic icons, independent of icon meaning, nor has an experiment permitted participants to give haptic icons their own descriptive names to aid in discrimination and learning; this premise forms the basis of our experiment.

3. LABORATORY EXPERIMENT

To explore our hypothesized Haptic Crayola Effect, we conducted a laboratory experiment. The experiment tested participants' ability to distinguish and recall haptic icons under three conditions: the icons had no names, the icons had predetermined non-descriptive names, and the icons had self-selected names chosen by participants.

3.1 Tasks

The experimental tasks were divided into two major phases, the learning phase, in which participants gained familiarity with the set of icons and their names if applicable, and the testing phase, in which participants attempted to recall sequences of icons. The differences between the three conditions appeared only in the learning phase. Ten haptic icons were chosen randomly from the set designed by [27], and these ten icons were used by all participants for all of the tasks.

3.1.1 Learning phase

The learning phase was designed to give participants exposure to the set of icons so that they would be able to distinguish them during the testing phase as well as to help the participants learn the names of the icons if they were in an applicable condition group. The tasks were designed to give participants equal exposure to the icons regardless of which condition group they were in. The tasks in the learning phase differed slightly based on condition.

3.1.1.1 No icon names condition

To gain an initial understanding of what haptic icons are and to begin working on the ability to distinguish the ten icons in the experiment, participants first simply felt each icon in sequence by clicking on a button labeled "next" to feel each icon. This task was repeated one more time.

After feeling each icon sequentially two times, participants completed a series of sorting tasks as used in previous experiments such as [16, 17]. The intention of these tasks was to give participants more exposure to the icons and to help them understand the differences between the icons. Each participant completed six sorting tasks using each number of bins between two and seven once.

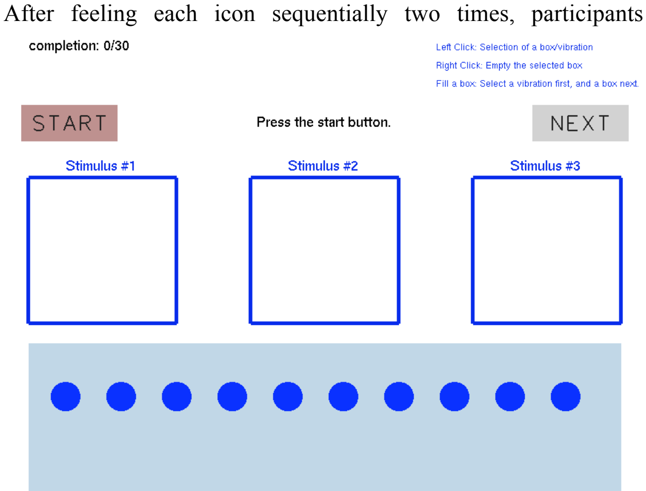

Figure 1: Screenshot of the testing phase application

3.1.1.2 Non-descriptive names condition

The only difference between this and the previously described condition is that beginning with the second time that the icons were felt in sequence, the pre-determined icon name was displayed on the computer screen each time an icon was played. The non-descriptive names were a subset of the menu labels used in an experiment by [9] and included nouns such as "hockey" and "leopard."

3.1.1.3 Self-selected names condition

This condition was the same as the previously described condition, except that participants chose their own names for the icons during the second time that the icons were felt in sequence. After each icon was played, the participant was prompted to type in a name. The participants were offered sheets of paper listing potential icon names: one sheet was a list of words that described emotions, one was words that described physical properties of objects, and one was words that described rhythms. Participants were told that the words on the sheets were suggestions, but they were allowed to use any names they desired. The only restriction was that no two icons could have the same name.

3.1.2 Testing phase

The testing phase was designed to determine participants' ability to recall sequences of icons. This phase was identical for all three conditions.

A screenshot of the interface used during the testing phase is shown in Figure 1. A trial began when the participant clicked the "start" button. At this point, each of the three target stimuli was played one time, with the corresponding box on the interface highlighting. There was a 1.5 second delay between each target stimulus. After feeling the target stimuli, participants attempted to find the targets from the full set of ten icons, which were represented as circles near the bottom of the screen and were presented in random order for each trial. Participants could feel each of the icons from the ten as many times as they needed to,

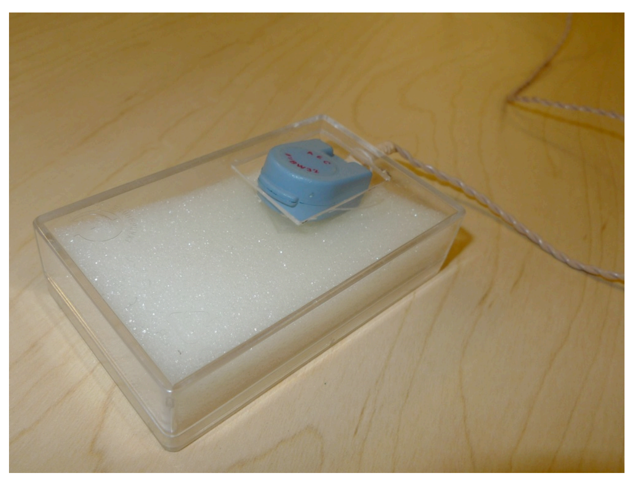

Figure 2: Experimental apparatus

but the target stimuli were only displayed the one time at the beginning of the trial. When a participant found the desired icon, they moved it to the box representing that target.

Each participant completed a total of 30 trials, with a one-minute break after every 10 trials.

3.2 Apparatus

The haptic icons were displayed using a Tactaid VBW32 voice coil actuator installed in an apparatus consisting of a layer of foam inside of a plastic box. This setup is shown in Figure 2. Participants' left index finger rested on the device during the experiment. The software was programmed in Microsoft Visual C++ with the OpenGL graphics library and ran on a 12.1" IBM X41 laptop with an optical mouse attached. The actuator was connected to the laptop's headphone jack, and the amplitude was maximized by using the highest volume settings on the laptop.

3.3 Participants

60 participants (32 female; age range from 19-44 with median 25) were recruited through email lists and posters on campus. 20 participants were randomly assigned to each of the three groups. Participants were paid $10 for their time. To motivate accurate performance, an additional $5 prize was awarded to the top 25% of participants.

3.4 Procedure

Each participant completed a single one hour session. After being introduced to the concept of haptic icons and to our apparatus, each participant completed the learning phase, the testing phase, and then a questionnaire.

The questionnaire asked for basic background information on the participant and then asked questions about their experience during the experiment based on their condition group. Portions of the questionnaire asked whether participants were able to remember icon names, whether they felt that the names were helpful, and whether they used any memory "tricks" in addition to the names.

3.5 Dependent measures

Accuracy was measured as the number of targets correctly identified out of the three for each trial.

Ability to remember icon names was self-reported in the questionnaire.

Completion time was calculated from the time that the final target stimulus was presented until the participant clicked on the "next" button in each trial.

3.6 Hypotheses

H1: Accuracy will be higher in the self-selected names condition than the other two conditions.

H2: Ability to remember names will be higher in the self-selected names condition than the non-descriptive names condition.

H3: Participants who report being able to remember the names better will have higher accuracy.

3.7 Design

The experiment used a 3 (conditions) x 30 (trials) mixed factor design with conditions being a between subjects factor and trials being a within subjects factor.

4. RESULTS

All participants reported at the least an intermediate level of computer use proficiency. An overwhelming majority of the participants claimed to use a mobile device which makes use of vibratory notification signals, such as a mobile phone or PDA. Of this majority, 29 were aware of or made use distinct vibratory signals for different notifications, such as a unique signal for an incoming text message and another for an incoming call. A small portion of participants were familiar with vibratory signals or force feedback used in video game controllers.

4.1 Matching task accuracy

On average, over the course of 30 trials, participants in the self-selected names group correctly matched 1.668 of the three targets; those in the non-descriptive names group correctly matched 1.515 of the three target icons; and those in the no-names group correctly matched 1.570 of the three target icons (Std. Error = 0.096). These means are shown in Figure 3. A 3 (condition) x 30 (trials, repeated measures) mixed ANOVA failed to show a significant difference for condition (between subjects $p = 0.524$, $F = 0.654$), but did show a significant difference for trial (within subjects $p < 0.001$, $F = 815.993$, partial $\eta^2 = 0.935$); post-hoc comparisons between trial mean accuracy scores across the three conditions using the Bonferroni adjustment indicated many significant differences between trial pairs, however a general trend in accuracy performance was not apparent. Overall mean accuracy scores for all participants from trials #11 and #22 were typically higher than the mean scores of most other trials; it is worth noting that these trials were performed immediately or shortly after participants took 1-minute breaks after trial #10 and #20.

4.2 Matching task completion time

On average, over the course of 30 trials, participants in the self-designated names group completed a matching task in 33.077s; those in the non-descriptive names group completed a matching task in 31.993s; and those in the no-names group completed a matching task in 34.995s (Std. Error = 2.380). A 3 (condition) x 30 (trials, repeated measures) mixed ANOVA failed to show a significant difference for condition (between subjects $p = 0.667$, $F = 0.408$), but did show a significant difference for trial (within subjects $p < 0.001$, $F = 589.407$, partial $\eta^2 = 0.912$); post-hoc

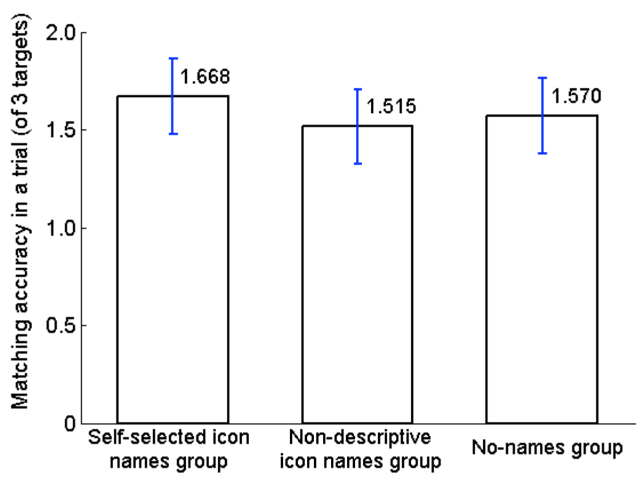

Figure 3: Matching task accuracy (N=60).

comparisons between trial mean completion times across the three conditions using the Bonferroni adjustment indicated many significant differences between trial pairs, with the matching task in the first trial tending to take significantly longer to complete than the those in the remaining trials.

4.3 Memory of icon names

The post-experiment questionnaire included a statement regarding whether the icon names used in the self-designated names group or in the non-descriptive names group were remembered at the end of the experiment. Subjects responded to this statement on a Likert scale ranging from *strongly agree* to *strongly disagree*. Of those in the self-designated names group, 8 participants reported to have remembered the names which they gave to the ten icons at the end of the experiment (answering *agree* or *strongly agree*), 6 reported to have forgotten the icon names (answering *disagree* or *strongly disagree*), and 6 remained neutral with respect to the statement. Of those in the non-descriptive names group, 3 participants reported to have remembered the names were given to the 10 icons, 12 reported to have forgotten the names, and 5 remained neutral. We compared accuracy and completion time for the matching task between the 11 participants who claimed to remember the icon names with the the 18 participants who claimed to have forgotten the icon names. A non-parametric Mann-Whitney test was conducted to compare the overall accuracy, or total number of correct matchings over 30 trials (90 matchings), for the group who remembered the icon names (N = 11, M = 52.55, SD = 11.46, mean rank = 17.45) with the group that forgot the icon names (N = 18, M = 46.33, SD = 15.33, mean rank = 13.50), however no significant difference was found between the groups (p = 0.225, U = 72.00). A second Mann-Whitney test was conducted to compare the average total matching task completion time over 30 trials for the group who remembered the icon names (N = 11, M = 1025.46s, SD = 325.59s, mean rank = 16.36) with the group that forgot the icon names (N = 18, M = 969.12s, SD = 265.45s, mean rank = 14.17), however no significant difference was found between the groups (p = 0.500, U = 84.00).

4.4 Self-selected icon names

We recorded the names given to the ten haptic icons by the 20 participants in the self-selected names group, for a total of 200 icon names. Over half of these names were purely rhythmic descriptions, defined in terms of the number of short and long pulses in the icon; a typical name would be "one short, three quick". The next largest group of words used were additional rhythm descriptors, often used in combination with the number of pulses; these include *scattered*, *elongated*, *punctuated*, and *staccato*. 12 of the 20 participants in this group used solely rhythmic descriptions, this included one participant who named most of the icons using a coding strategy of small and lower case 'B's, presumably to indicate long and short tones; an example being *bBbb*. 6 users used a combination of rhythmic and other words, with rhythm-based names forming the majority. Other words included emotional adjectives, such as *angry*, *impatient*, or *talkative*, as well as actions or objects in the real world, such as *hop & skip*, *warning*, *mosquito*, *inkjet*, *butler*, and *doorbell*. One participant used solely emotional words. The remaining participant in this group used solely real-world objects as names.

4.5 Questionnaire results

The post-experiment questionnaire queried users in each group about their ability to distinguish and remember the icons throughout the experiment. Participants were asked to describe any tricks or mnemonics used to aid in either icon discrimination or learning, aside from the names.

Many participants in the self-selected names group reported to use vibration spacing or pause length as additional memory cues for icon discrimination and learning. After the learning phase, seven participants in this group reported to have no confidence in their ability to associate the icons with the names they gave to the icons; nine participants remained confident in their ability to recall this association; the remaining 4 were uncertain of their ability. Ten participants in this group claimed that the icon names were effective for helping to discriminating the icons during the testing phase; nine claimed that the names were ineffective. Only seven participants in this group claimed that the names they gave to the icons accurately described the icons; five participants did not share this sentiment, while the remaining eight participants were undecided. The majority (N = 12) of participants in this group claimed that more descriptive names for the icons would have helped in discriminating and remembering the icons in the testing phase.

In the non-descriptive icons names group, participants mentioned counting and distinguishing between short and long pulses. Other tactics included foot-tapping along with the vibrations, associating sounds to the vibrations, visualizing the series of vibrations spatially as "peaks on a graph", or by attempting to map the vibrations to Morse code. Six participants in this group did not feel confident in their ability to recall the name-stimulus association after the learning phase, while nine did feel confident. The majority (N = 11) of participants in this group did not feel as though the non-descriptive icon names helped them distinguish or remember the icons during the testing phase, and felt as though more descriptive icon names would have improved their performance (N = 12), and that they would have selected better names for the icons (N = 13), had they been given the opportunity.

Participants in the no-names group reported using a similar variety of tricks or mnemonics for distinguishing and and learning the icons throughout the experiment; many paid attention to the number and spacing of long and short pulses or pauses. Others attempted to visualize the sequence as the lines and dots of Morse code, moved their heads along with the vibration, hummed along with the sequence, or counted the number of pulses on the fingers of their free hand. The majority of participants in this group (N = 13) agreed that giving the icons names would have assisted them in the process of distinguishing and remembering the icons.

4.6 Results summary

H1: Accuracy will be higher in the self-selected names condition than the other two conditions. *Not supported.*

H2: Ability to remember names will be higher in the self-selected names condition than the non-descriptive names condition. *Supported by self-reported questionnaire responses.*

H3: Participants who report being able to remember the names better will have higher accuracy. *Not supported.*

5. DISCUSSION

From the post-experiment questionnaire, we see that the use of names in our experiment was not as beneficial as it could have been. In the no-names condition, 70% of participants agreed with the statement, "Giving the icons names would have helped me remember them and tell them apart." Despite this strong desire for these participants to have icon names, there was no significant difference between the different naming conditions. This suggests that our particular method of providing and learning icon names was not as effective as possible.

This suspicion is supported by participants' self-reported lack of learning. In the questionnaire, only 27.5% of participants agreed that they could remember icon names (40% in the self-selected names condition and 15% in the non-descriptive names condition). As expected, the self-selected names group reported a higher rate of remembering the icon names, but their figure is still disappointingly low. Because these values rely on self-report, we cannot be sure that the numbers are accurate, but there is no obvious reason for participants to have responded dishonestly.

Furthermore, 60% of participants in each of the self-selected names condition and the non-descriptive names condition agreed to a statement, "More descriptive names would have improved performance in the matching task." This suggests that the names in both conditions were less than ideal.

We suspect that more appropriate names and a better memorization of names may have resulted in more difference between the condition groups. In the remainder of this section, we discuss the names used in our experiment, examine the limitations of our learning phase, and suggest alternative ways to help differentiate haptic icons.

5.1 Verbal abstraction of stimuli

5.1.1 Trends in self-selected icon names

We examine the icon names chosen by participants. Since the ten vibration stimuli in this study differed only by their simple rhythms, most of names that participants chose included only descriptions of the rhythm. Names that involved emotion or physical characteristics were very rare in the experiment. We speculate that this kind of rhythmic description results in worse performance than abstract, yet still descriptive, names would have.

We have two explanations for participants' choice of rhythmic descriptions rather than abstract names. First, as discussed in the related work section, Gabrielsson suggested that people primarily use structural rhythmic properties in adjective rating of auditory stimuli [10]. Since vibratory sensation is very similar to auditory sensation, we can expect similar results here.

The second potential reason is related to the level of verbal abstraction. In the learning phase, participants were asked to name each icon on their second exposure to it. Therefore, the verbal abstraction of vibratory sensation was at a primitive level and direct perceptive descriptions were used rather than emotional or cognitive metaphors to represent a vibratory stimulus.

5.1.2 Verbal overshadowing

Our results may be related to Schooler and Engstler-Schooler's concept of verbal overshadowing [22] as discussed in the related work section. When people make verbal description of a complex stimulus, their identification performance on the stimulus can be lower than without the verbal description. We suspect that since we used relatively simple stimuli, the effect of verbal overshadowing was small if present at all, but it is possible that this effect led to reduced performance in the two conditions with icon names.

5.2 Learning task design

The tasks in our learning phase were not ideal for learning either the set of icons or their names. MacLean [18] recommends the use of practical, regular-use scenarios for learning icons, and the inclusion of reinforced learning when possible, providing feedback to the user of their progress in terms of learning the icons. MacLean also points out that long, stressful experimental sessions are likely an ineffective way to learn. Work from Enriquez suggests that it may be beneficial to allow the user to return to an exploratory learning phase if the performance results of a reinforced learning phase are too low [7].

Our primary concern when designing our learning phase was that exposure to the icon sets remains equal across the three conditions. Therefore, we could not include any tasks that related specifically to the names or tested the learning of the names, because these tasks would not have any equivalent for the group with no icon names.

Our learning phase was also relatively short for the sufficient memorization of each vibration stimulus and its name. Many participants remained unfamiliar with various vibration patterns and their discrimination. Future work in this area should include a longitudinal study with enough time and experience to fully memorize the icons and their names.

5.3 Hierarchy in naming

Many object names have a clear hierarchical structure. An example is binomial nomenclatures of plants and animals, where each item has a clear place within a hierarchical structure. For example, we can intuitively notice *Panthera tigris altaica* (Siberian tiger) is close to the *Panthera leo* (lion) and *Panthera onca* (jaguar), and even closer to the *Panthera tigris jacksoni* (Malay tiger).

We suspect that a clear hierarchical grouping would be helpful in identifying and memorizing the characteristics and name of a haptic stimulus. This is further supported by Blattner's [1] work with hierarchical auditory icons discussed in the related work section.

Another piece of potential future work would be to design a set of hierarchical icon structures and names to examine what effect that would have on learning, memorization, and recall.

5.4 Descriptive non-verbal representation

This study examined descriptiveness and verbal representation with the three conditions being non-descriptive verbal names, descriptive verbal names, and no names. Another potential condition would have been descriptive non-verbal names. Representation of the stimuli using graphical symbols could be introduced for this condition in future work. In this case, the graphic symbol should be selected carefully to avoid the direct translation of symbol into a verbal representation in an early cognition stage. For example, waveform amplitude envelope of a vibration stimulus could be a good choice of symbol.

6. CONCLUSIONS

Our experiment suggests that the naming of haptic icons, whether by using non-descriptive abstract names or descriptive user-generated rhythm-based names, has no effect on user performance in a haptic icon matching task requiring differentiation and memory. Also, participants who claim to able to remember the names do not perform better in terms of accuracy. Participants in the self-selected names group reported to remember the icon names to a greater degree than those in the non-descriptive names group, however this did not lead to increased accuracy performance.

As described in the discussion section, we believe that the lack of difference in performance between our groups may have been related to the types of icon names used and a non-optimal learning phase in the experiment. Also possible is the detrimental role of verbal overshadowing: the inability to remember complex non-verbal stimuli using verbal memory. Future work should include similar experiments exploring the effects of different types of naming including hierarchical names and non-verbal names while including a more focused learning phase to ensure that the names are memorized. Until this time the Haptic Crayola Effect remains to be seen.

7. REFERENCES

[1] Blattner MM, Sumikawa DA, Greenberg RM. Earcons and icons: Their structure and common design principles. Human-Computer Interaction. 1989;4(1):11-44.

[2] Brewster SA, Chohan F, Brown LM. Tactile feedback for mobile interactions. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07. 2007:159.

[3] Brown LM, Brewster SA. A first investigation into the effectiveness of tactons. Proc. World Haptics. 2005:167-176.

[4] Chan A, MacLean KE, McGrenere J. Learning and Identifying Haptic Icons under Workload. First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. 2004:432-439.

[5] Chan A, MacLean KE, McGrenere J. Designing haptic icons to support collaborative turn-taking. International Journal of Human-Computer Studies. 2008;66(5):333-355.

[6] Crayola LLC. *Crayola Fun Facts*. Available at: http://crayola.com/mediacenter/download/news/press_release_81.pdf

[7] Enriquez MJ, MacLean KE, Chita C. Haptic phonemes. Proceedings of the 8th international conference on Multimodal interfaces - ICMI '06. 2006:302.

[8] Enriquez MJ, MacLean KE. The Role of Choice in Longitudinal Recall of Meaningful Tactile Signals. 2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. 2008:49-56.

[9] Findlater L, Moffatt K, McGrenere J, Dawson J. Ephemeral Adaptation: The Use of Gradual Onset to Improve Menu Selection Performance. Proceedings of CHI 2009. 2009:1655-1664.

[10] Gabrielsson A. Adjective ratings and dimension analyses of auditory rhythm patterns. Scandinavian Journal of Psychology. 1973;14(4):244-260.

[11] Gaver WW. Auditory Icons: Using Sound in Computer Interfaces. Human-Computer Interaction. 1986;2(2):167-177.

[12] Howard DM, Angus J (eds.). *Acoustics and psychoacoustics* (3rd edn). 2006. Focal Press, Oxford UK.

[13] Ishii R, O'Mahony M. Taste sorting and naming: can taste concepts be misrepresented by traditional psychophysical labelling systems? Chemical Senses. 1987;12(1):37-51.

[14] Klatzky RL, Lederman SJ. Tactile roughness perception with a rigid link interposed between skin and surface. Perception & psychophysics. 1999;61(4):591-607.

[15] Lehrner JP, Glück J, Laska M. Odor identification, consistency of label use, olfactory threshold and their relationships to odor memory over the human lifespan. Chemical senses. 1999;24(3):337-46.

[16] Luk J, Pasquero J, Little S, Maclean KE, Levesque V, Hayward V. A role for haptics in mobile interaction: initial design using a handheld tactile display prototype. In: Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM; 2006:180.

[17] MacLean KE, Enriquez MJ. Perceptual design of haptic icons. In: Proceedings of Eurohaptics. Citeseer; 2003:351-363.

[18] MacLean KE, Hayward V. Do It Yourself Haptics: Part II [Tutorial]. IEEE Robotics & Automation Magazine. 2008;15(1):104-119.

[19] Melcher JM, Schooler JW. The Misremembrance of Wines Past: Verbal and Perceptual Expertise Differentially Mediate Verbal Overshadowing of Taste Memory. Journal of Memory and Language. 1996;35(2):231-245.

[20] Parr WV, Heatherbell D, White KG. Demystifying wine expertise: olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. Chemical senses. 2002;27(8):747-55.

[21] Sandhofer CM, Smith LB. Learning Color Words Involves Learning a System of Mappings. Developmental Psychology. 1999;35(3):668-679.

[22] Schooler JW, Engstler-Schooler TY. Verbal overshadowing of visual memories: some things are better left unsaid. Cognitive psychology. 1990;22(1):36-71.

[23] Swerdfeger BA. A First and Second Longitudinal Study of Haptic Icon Learnability. M Sc thesis. 2009.

[24] Swerdfeger BA, Fernquist J, Hazelton TW, MacLean KE. Exploring melodic variance in rhythmic haptic stimulus design. Proceedings of Graphics Interface 2009. 2009:133-140.

[25] Tan HZ, Srinivasan MA, Eberman B, Cheng B. Human factors for the design of force-reflecting haptic interfaces. Dynamic Systems and Control. 1994;55(1):353-359.

[26] Tan HZ, Durlach NI, Reed CM, Rabinowitz WM. Information transmission with a multifinger tactual display. Perception & psychophysics. 1999;61(6):993-1008.

[27] Ternes D, MacLean KE. Designing large sets of haptic icons with rhythm. Lecture Notes in Computer Science. 2008;5024:199.

[28] Tufte ER. *Envisioning Information*. Graphics Press LLC, Cheshire Connecticut. 1990.

[29] van Erp JB, Spapé MM. Distilling the underlying dimensions of tactile melodies. In: Proceedings of Eurohaptics.Vol 2003.; 2003:111-120.

[30] Ware C. *Information Visualization: Perception for Design*. Morgan Kaufman Publishers, San Fransisco. 2004; 112-122.

[31] Weiser M, Brown J. Designing calm technology. Powergrid Journal. 1996;01(01):94-110.