

# Towards the Non-Visual Monitoring of Canine Physiology in Real-Time by Blind Handlers

Sean Mealin<sup>1</sup>, Mike Winters<sup>1</sup>, Ignacio X. Domínguez<sup>1</sup>, Michelle Marrero-García<sup>2</sup>, Alper Bozkurt<sup>3</sup>, Barbara L. Sherman<sup>4</sup>, David L. Roberts<sup>1</sup>

<sup>1</sup> Department of Computer Science

<sup>2</sup> CROEM Specialized Residential School

<sup>3</sup> Department of Electrical and Computer Engineering

<sup>4</sup> College of Veterinary Medicine

North Carolina State University

spmealin@ncsu.edu

## ABSTRACT

One of the challenges to working with canines is that whereas humans are primarily vocal communicators, canines are primarily postural and behavioral communicators. It can take years to gain some level of proficiency at reading canine body language, even under the best of circumstances. In the case of guide dogs and visually-impaired handlers, this task is even more difficult. Luckily, new technology designed to help monitor canines may prove useful in helping handlers, especially those with visual impairments, to better understand and interpret what their working partners are feeling or saying.

In prior work a light-weight, wearable, wireless physiological monitoring system was shown to be accurate for measuring canines' heart and respiratory rates [6]. In this paper, we consider the complementary problem of communicating physiological information to handlers. We introduce two non-visual interfaces for monitoring a canine's heart and respiratory rates, an audio interface and a vibrotactile interface. We also present the results of two initial studies to evaluate the efficacy of the interfaces. In the first study we found that many participants were more confident in detecting changes in heart and respiratory rate using the audio interface, however most of the time they were just as accurate with the vibrotactile interface with only a slight increase in detection latency.

## INTRODUCTION

Humans and canines communicate information about themselves and their surroundings in different ways. Humans primarily use verbal communication, while canines typically use more visual means, such as body language. As words can have different meanings in different contexts, there are similar nuances to canine body language and behavior. Like learn-

ing a foreign language, it becomes easier to interpret the visual language of the canine with practice. Using technology to assist with translating body language to something more intuitive for people would ease and increase human-canine communication, for example by helping novice dog trainers understand their dogs better. This would also provide information that is unavailable to those who cannot see their dog, such as blind handlers and guide dogs.

When working with a guide dog, a blind handler is reliant on auditory clues and the physical feedback provided by the leash or harness to guess at the actions that the guide dog is taking; however, the dog may take the same action for many reasons. For example, if a dog suddenly moves forward, it could be because the dog has seen someone that they recognize and is trying to show the handler, or because the dog is lunging towards food, or because the dog is fearful of a car approaching the handler. Without being able to identify the stimulus directly, the next best thing would be to identify the emotions that the dog is experiencing just before, or during the time of the action. Many of the body language hints that sighted handlers use to identify the emotional context of a dog's actions are unobservable to blind handlers. Previous work focused on the development of portable technology to monitor canine physiology [6], which we can use to give a handler hints about a canine's emotional state.

This paper presents the design of two non-visual interfaces to allow the monitoring of a canine's heart rate and respiratory rate. The interfaces were designed to communicate physiological information to a blind handler in real-time, while being as discreet and non-distracting as possible. The interfaces must be discreet, since the handler may be working in a quiet area such as a library. They also must be non-distracting, since the handler may be doing something that requires her attention, such as crossing a busy street or conversing with a colleague. To meet these goals, we based the interfaces on two modalities: audio and vibrotactile feedback.

We will present the results of two preliminary studies performed using these two interfaces, using a total of 15 sighted participants and simulated data based on data collected by existing physiological monitoring hardware [6]. For the first

study, we asked participants to monitor the interfaces and give verbal feedback when there was a change in heart or respiratory rate. We found for almost all participants, the audio interface was preferred initially, however many said that they would find it very distracting and would eventually switch to the vibrotactile interface. Despite initial preferences, participants were more accurate using the latter interface during some tasks, with only a small increase in latency time. For the second study, we introduced a concentration-based task to evaluate the distraction level of the participants. We found that as the audio-based concentration task became more difficult, the latency for the audio interface increased while the latency for the vibrotactile interface decreased.

## RELATED WORK

The idea of using audio to communicate data to the blind and others is not a new concept. Line graphs, pie charts, and other types of plots have been presented to the blind using changes in pitch, timbre, and stereo with varying levels of success [11, 20]. Audio interfaces are becoming more common in vehicles [12, 15]. Node-link diagrams, a standard data construct in science, technology, engineering, and mathematics (STEM) disciplines, have been made accessible via synthesized speech [2]. Medical data has also been communicated through sonification, however that study focused on communicating brainwave patterns [1]. The real-time monitoring of financial data has also been studied using audio; however much of the information was coded using music theory, which could be difficult to interpret for the untrained user [10].

Vibrotactile interfaces have been likewise explored in previous studies; they have been used to provide directional and spatial cues [18], they have been used to augment audio communication [9], and have become a universal augmentation for mobile devices [17]. Vibrotactile interfaces have been studied in visually complex environments, like alerting pilots to changes in instrumentation in a plane's cockpit [19]. The development of Tactons, which are the vibrotactile equivalent of phonemes, have allowed the effect of intensity, amplitude, and frequency to be studied in various applications [3]. When applied to human physiological monitoring, a vibrotactile alert system proved effective both when used alone, and in tandem with other notification systems [16].

This work is based on previous efforts to introduce technology to the human-canine relationship. Past studies have focused on the development of wearable technology for dogs, with an emphasis on using accelerometers and gyroscopes to allow computers to identify canine posture and behavior. That information was then applied to automated training systems [7, 8]. This project also directly depends on work done to allow the noninvasive monitoring of canine physiology [6]. Using the two non-visual modalities outlined above, and the physiological monitoring systems, our work is the first to allow a completely blind user to monitor their dog's heart and respiratory rate in real-time while performing everyday tasks.

## PLATFORM

The focus of this paper is the handler's interface. Previously, it has been shown that a lightweight physiological monitoring

system is possible [6], so we do not address the methodology used to capture the heart and respiratory rate of canines in great detail. For the purposes of this paper it is sufficient to know that it is possible to monitor a canine's heart rate, heart rate variability, and respiratory rate outside of a laboratory environment using both electrocardiogram (ECG) and photoplethysmogram (PPG) techniques [6].

## Common Hardware

We based the two interfaces on a common platform and commercially available components, which we then extended with custom hardware as necessary. A BeagleBone Black, a microcomputer approximately the size of a deck of cards, serves as the central hub. The BeagleBone Black is equipped with an ARM Cortex-A8 CPU running at 1.0 gigahertz, 512 megabytes of memory, 4.0 gigabytes of onboard flash memory, and runs Ubuntu 14.04 as an operating system.

## Software

We connected the interfaces to a software system we wrote that mimics a canine's physiological status to make experiments consistent and reproducible. The system allows us to write virtual walks, which are broken down into separate behaviors such as running and walking. It then uses those behaviors to output events, which we define as a time and action pairing. The interfaces then read those events, and communicates the action, such as a heart beat or breath.

Each interface is capable of operating in two modes. Absolute mode communicates each heart beat and breath, while relative mode only indicates when the heart or respiratory rates change by a certain amount. The advantage of the first is that users can "tune in" to the interface whenever they wish to see their dog's status, while the latter generates fewer events. The disadvantage is that absolute mode is noisier, while a distracted user may lose context when using relative mode.

## Audio Interface

The audio interface consists of a bluetooth earpiece connected to the BeagleBone Black which is mounted on a PVC pipe assembly cut to the shape of a standard guide dog handle. Since the environmental sounds that blind people use for getting around are important, the earpiece is lightweight, very small, and is wearable on either ear. Future improvements to this interface could include using bone-conduction, which generates audio through vibrations directly applied to the bones on the side of the face, and does not block environmental sounds.

For communicating information to users via audio, a variety of techniques and best practices have been identified [5]; however, the most common techniques either masque environmental audio, or are unavailable given the mono-audio output of the earpiece. One solution is to use different volume levels to separate heart and respiratory rate; however, studies have found that users are not proficient at judging small to moderate differences in relative volume, and can become quickly annoyed with the perceived sound pollution [4].

The audio interface uses different short sounds to communicate to users. In absolute mode, a double-thud sound of a heart beating represents a heartbeat, while a clip of a dog



**Figure 1. The vibrotactile interface, annotated with the components (a). An illustration of a typical off-center grip (b).**

panting represents a breath. In relative mode, a high-pitched tone that increases in pitch represents heart rate getting faster, while a high pitch tone that decreases in pitch means that it is getting slower. For respiratory rate, the semantic meanings are the same, but the pitch of the tones is significantly lower. We chose the pitch of the sounds to fall within the range of 150 Hz to 5,000 Hz, which has been shown to be effective for communicating information to users [5].

### Vibrotactile Interface

We based the shape of the vibrotactile interface on a guide dog handle, and used the data from anonymous pictures of how guide dog users grip the handle when working with their dog as a guide on where to position the ERM motors. We mounted The BeagleBone Black and a battery on a cross-pipe, while we embedded the ERM motors into either side of the grip, which we then wrapped in common grip tape to make it more pleasant to hold. Figure 1(a) gives an overview of the handle, indicating the BeagleBone Black, grip, and ERM motors. Figure 1(b) shows the interface being held.

The two motors allow us to communicate heart rate through one side of the handle, and respiratory rate through the other. The PVC piping is rigid enough that the vibrations are easily felt at any point on the grip, but are dampened enough that depending on the location of the hand, one source of vibration is more noticeable than the other. A majority of guide dog users grasp the handle closer to the right side, so the channel that communicates heart rate was placed there and slightly reduced in intensity compared to the left motor which communicates respiratory rate.

High frequency vibrations tend to get an observer’s attention more quickly, but can become irritating and be ignored when left active for long durations [13]. Consequently, the vibrations for this interface were designed to be high frequency, but as short as possible in duration. When in absolute mode, the

interface communicates each heartbeat by a short pulse, followed by an even shorter pause, and lastly a vibration pulse of equal duration to the first, which mimics the common “thud-thud” representation of the heart. Breath is represented as a soft vibration that increases in intensity to a peak, followed by a decrease to stillness. While not as intuitive as heart rate, we hoped users would associate this with an inhale of breath, which gradually gets louder before tapering off in volume. When in relative mode, a double-pulse pattern represents an increase in heart rate, while a rapid series of eight pulses represents a decrease. For respiratory rate, an increase is indicated by a soft vibration that quickly and smoothly increases before abruptly stopping, while a decrease in respiratory rate is an intense vibration that smoothly decreases to stillness.

### METHODOLOGY FOR STUDY ONE

For the first study, to evaluate the efficacy of the two interfaces, we had participants go on a series of virtual walks. During each experiment, which lasted on average 1.5 hours, participants had the opportunity to take breaks after each simulated walk to minimize mental fatigue.

The 10 participants were recruited via word of mouth from around the Computer Science department. The only disqualifying criteria was prior experience using the interfaces during design time. Six participants were graduate Computer Science students, while the other four were from other departments, or nonstudents. All of the participants were sighted, and volunteered for the study for no compensation.

### Experimental Design

Participants used the interfaces and modes in random order. To reduce the burden of switching between interfaces, we forced the modes to be contiguous; for example, participants completed both modes of the vibrotactile interface before the audio interface. During each interface-mode combination, the participant went on a total of six virtual walks.

Space prevents a detailed description of all the walks. Therefore, as a representative example, we present walk *W1* here, which was always used as a training walk so the participant could see how the interface worked. The walk started at heart rate = 75 and respiratory rate = 15 for 15 seconds, before increasing to  $HR = 95$  and  $RR = 25$  over a 10 second period. Both remained constant for 20 seconds, before decreasing back to  $HR = 75$  and  $RR = 15$  over 10 seconds, where it remained for the last 15 seconds of the walk. Since this walk was used to show the participant both a simple increase and decrease in heart and respiratory rate, the walk was explained to them ahead of time. All of the remaining walks were defined in a similar manner.

### Participant Experience

We gave participants a general overview of the interfaces, and explained that they would be monitoring heart and respiratory rate of a simulated canine, while giving verbal indications when the rates changed. We then told them which interface-mode combination they would be using first.

There were seven steps for each interface-mode combination.

1. **Training:** We showed the participant how the interface-mode combination represented heart and respiratory rate. For example, the training prompt for absolute mode was: “The [sound/vibration] you are about to [hear/feel] represents the dog’s [heart beat/breath].” The participant could repeat training before progressing.
2. The second step consisted of the participant going on walk W1, which was the training walk.
- 3-7. 5 walks: For each walk, the pre-walk prompt was “for this walk, the heart rate will start at [HR starting value], while the respiratory rate will start at [RR starting value]. Please verbally indicate if you think that the heart or respiratory rate of the dog is increasing or decreasing. Do you have any questions before we begin?” If the participant said no, the walk would then start. After the walk, we asked each participant two questions. The first question was “on a scale from one to ten, how confident are you that you were able to detect increases and decreases of heart and respiratory rate?” while the second question was “Do you have any comments about this walk?”

After each interface-mode combination, we asked the participant if they had any comments that they would like to share.

## RESULTS OF STUDY ONE

To evaluate the interface-mode combinations, we focused on three key metrics: accuracy, latency, and confidence. To compare the first two metrics between conditions (interface and mode), we conducted an ANOVA. For every metric, a significant factor was the simulated walk; however, we intentionally designed each one to have a varying difficulty level, so we aggregate that factor in the analysis below.

### Accuracy

For each behavior, we compared the participant-generated indications to the ground truth. We totaled the number of times that the participant correctly identified an increase or decrease (true positive), and correctly gave no indication when the heart and respiratory rates were steady (a true negative). Note that two participants were unable to hear the audio signal for changes in respiratory rate when using the audio interface in relative mode, however, that data is included in the analysis to give a complete picture of the interfaces.

For the accuracy of the heart rate detection task, the only significant factor was mode ( $F = 48.068$ , d.f. = 1,  $p < 0.001$ ). When operating in absolute mode, the mean accuracy was 81.52% with a variance of 18.82. When in relative mode, the mean was 96.20% with variance 13.09.

Likewise for the respiratory rate accuracy, the significant factors were mode ( $F = 31.158$ , d.f. = 1,  $p < 0.001$ ) and interface ( $F = 7.079$ , d.f. = 1,  $p = 0.009$ ). In absolute mode, the mean and variance were 75.67% and 19.86 respectively, while in relative mode, the mean was 90.60% with variance 19.81. When participants used the audio interface, the mean accuracy was 79.57% with variance 23.20, while the vibrotactile interface resulted in mean 86.69% and variance 18.32.

### Latency

We defined latency as the difference in time between the first heart beat or breath during an increase or decrease and the time that the participant verbally stated that there was a change. If the participant did not identify a change in either rate, we excluded that particular behavior from the calculation for that participant.

For latency during the heart rate detection task, both mode ( $F = 402.140$ , d.f. = 1,  $p < 0.001$ ) and interface ( $F = 4.263$ , d.f. = 1,  $p = 0.04$ ) were significant. For absolute mode, the mean latency was 13.89 seconds, with variance 5.75, while relative mode scored mean latency 4.84 seconds, with variance 1.42. When the audio interface was used, the mean and variance were 8.94 and 6.04 respectively, while the vibrotactile interface resulted in mean latency 9.78, with variance 6.29.

For the latency during the respiratory rate detection task, the only statistically significant factor was mode ( $F = 233.644$ , d.f. = 1,  $p < 0.001$ ). In absolute mode, the mean latency was 15.63 seconds, with variance 7.56. In relative mode, the mean and variance was 5.34 and 1.83 respectively.

### Confidence

After each walk, we had the participants rate how confident they felt that they detected the changes in heart and respiratory rate on a scale from one to 10, where one was not confident at all and 10 was very confident. We used these results in combination with comments that participants made to get an idea of how easy they felt the interfaces were to use.

Participants felt the most confident with the audio interface in relative mode, assigning it an average score of 8.92. The vibrotactile interface in the same mode received the next highest score, with participants giving it 8.53 on average. There was the same pattern for the audio and vibrotactile interfaces in absolute mode, with average confidence scores of 6.4 and 5.96 respectively.

## DISCUSSION OF STUDY ONE

### Accuracy

For the accuracy during the heart rate detection task, both modes scored very well, with accuracies greater than 80%. Participants using the relative mode scored a higher mean accuracy by approximately 15%, which indicates that it was easier to detect changes in relative mode. A possible explanation of this result is that users found it easier when they did not have to keep track of the heartbeats themselves.

Just as the heart rate accuracy task, mode was significant, and the relative mode proved to be more accurate than the absolute mode. Both scores were lower than their counterparts during the heart rate detection, which could indicate that users focused more on the heart rate; however, we do not have enough data to be sure. Unlike the heart rate task, the interfaces themselves were significant, with the vibrotactile interface doing moderately better than the audio interface. This is interesting considering that many participants felt more confident with the audio interface, as explained below. As previously mentioned, two participants were seemingly unable to

detect the audio indicators for changes in respiratory rate in relative mode. When their data is removed from the analysis, the accuracy of both interfaces increase by approximately 5%, however interface is no longer a statistically significant factor ( $p > 0.05$ ). Considering that participants performed this study in a quiet environment, this indicates that we need to reevaluate the frequency of the audio signal; however, this could also indicate that the audio interface is more susceptible to interference in noisy environments.

### Latency

For the latency during the heart rate detection task, both interface and mode were significant. Surprisingly, the audio interface resulted in slightly lower latencies than the vibrotactile interface, which could indicate that users are more accustomed to using their auditory sense rather than their tactile sense for detecting alerts. For both of these interfaces, there was a relatively high variance, which indicates that some participants were able to do better than others. When examining relative mode compared to absolute mode, relative mode resulted in significantly lower latencies, which we expected, since that mode only indicates when something changes.

For the respiratory rate latency, once again the only significant factor was the mode. For the same reason as previously outlined, we were not surprised to find that relative mode did significantly better than absolute mode. Since responding to a change in a canine's physiological status is time critical, it appears that relative mode may be more useful due to the lower latency times.

### Confidence

Despite feedback enabling a ranking for the interface-mode combinations as a whole, subjectively, participants could not agree which interface was best. All of the participants felt more confident using relative mode; however five participants indicated that the audio interface was their favorite ("very easy to identify what's going on", and "just tells you when something is happening"), while four participants said they were very distracted by the audio ("it sounds like an alarm, like something urgently needs my attention"), and would find it difficult to navigate safely while using the interface. The audio interface in absolute mode was more distracting, with six participants saying they would be unable to concentrate on anything else ("the constant sound is very distracting, there's no way I could talk to someone with this on"). Interestingly, despite 80% of participants feeling less confident with the vibrotactile interface in either mode, results were as accurate as the audio interface. The two participants that preferred the vibrotactile interface self-identified as getting distracted by audio ("I'm just not an audio person. I even get distracted any time there's stuff like music."). Four participants said that they would eventually prefer the vibrotactile interface ("After a few months, I would really like it").

## METHODOLOGY OF STUDY TWO

We introduced a distraction task in the second study. Like study one, participants went on a series of virtual walks. Each session lasted on average 30 minutes, and we had 10 participants (five of which had taken part in study one). We recruited

all participants by word of mouth, and none of them were involved with the design of the interfaces.

### Experimental Design

As with the previous study, participants went on virtual walks while verbally indicating changes in heart and respiratory rate for the four interface-mode combinations. There were a total of eight walks, which varied the heart and respiratory rate independently. Each walk lasted 60 seconds, and was designed using the same software as described above; however, we relaxed the requirement of modeling it off observed canine physiological data to add additional variation. In the interest of space, we are omitting the description of the walks.

During each virtual walk, we asked participants to perform an audio-based N-back test to place them under additional cognitive load while using the interfaces [14]. In an N-back test, a series of elements are sequentially communicated to the participant, who must indicate if the current element is the same as, or different than the element that occurred  $N$  elements ago. For our N-back test, the elements consisted of a computer-generated voice speaking the numbers "one," "two," and "three" at 2.5 second intervals. The participant indicated if the element was the same or different by pressing keys on a laptop keyboard.

For the first four walks, participants used all four interface-mode combinations while performing an N-back test with  $N = 1$ , while the second four walks had  $N = 2$ . A computer randomized the order of the walks and interface-mode combinations, with the restriction that all combinations were used exactly once for each value of  $N$ .

### Participant Experience

We gave each participant the same general overview of the project as we gave in the first study. We also explained what an N-back test was, and discussed how to indicate if an element was the same or different.

Without using the audio or vibrotactile interface, participants did a 60 second N-back test with  $N = 1$  to allow them to become accustomed to N-back tests. We then went through the following three steps for each interface-mode combination:

1. Training: We used the same prompts as we used for the training step in the first study.
2. Walks: Participants went on a virtual walk, while performing the N-back test. The prompt for this step was "Now you are going to go on a virtual walk with this interface. Please remember to verbalize when you think there is a change in either heart rate or respiratory rate. Please also do the N-back test as accurately and quickly as possible."
3. We then asked the participant if they had any feedback.

After participants used all four interfaces, they were given an opportunity for a five minute break. They then did a 60 second N-back test with  $N = 2$ . We then went through the above steps again, with the N-back test with  $N = 2$ . Participants could skip the training step if they wished.

## RESULTS OF STUDY TWO

Just as in the first study, we focused on accuracy and latency as metrics of performance. We once again derived the ground truth using the scripts of the simulated walks. The formulas and methodologies used to calculate the accuracy scores and latency times for heart and respiratory rate detection were identical to those used before.

New to this study was the accuracy and latency of the N-back test. Latency was defined as the amount of time between the start of the voice speaking the element, and the participant pressing a key. If the participant did not respond to an element within the 2.5 second window, the software counted that as the wrong answer and included the number of missed elements in the results of each test. To compare the results across conditions (interface, mode, and N-back) we conducted an ANOVA and report the results below.

### Accuracy

For heart rate detection, the only significant differentiator was the interface mode ( $F = 13.702$ , d.f. = 1,  $p = 0.0004$ ). When in absolute mode, the mean accuracy was 72.58%, with a variance of 30.88. When in relative mode, the accuracy was 92.42%, with variance 13.26.

Mode was likewise the only differentiator for the respiratory rate detection task ( $F = 22.167$ , d.f. = 1,  $p = 0.00001$ ). While in absolute mode, the mean accuracy was 62.75% with a variance of 28.14. When in relative mode, the mean accuracy was 88.38% with a variance of 17.52.

For the accuracy of the N-back tests, the only significant differences were between the levels of  $N$  ( $F = 4.538$ , d.f. = 1,  $p = 0.037$ ). For the interface-mode combinations during N-back tests with  $N = 1$ , the N-back accuracy was 73.04% with variance 16.55. For the walks with  $N = 2$ , the mean accuracy of the N-back tests was 65.24% with variance 15.91.

### Latency

There were multiple statistically significant results for the latency during the heart rate detection task. Mode was one differentiator ( $F = 26.519$ , d.f. = 1,  $p = 0.000003$ ), with absolute mode resulting in mean latency 12.864 seconds with variance 5.95, while relative mode resulted in mean latency 6.886 seconds with variance 3.85. Another significant result was the interaction between interface and N-back test level ( $F = 5.760$ , d.f. = 1,  $p = 0.020$ ). The audio interface resulted in mean latency 7.867 and variance 4.20 when  $N = 1$ , and latency 10.042 and variance 6.93 when  $N = 2$ . The vibrotactile interface resulted in mean latency 12.176 seconds and variance 6.75 when  $N = 1$ , and latency 9.125 and variance 4.03 when  $N = 2$ . Table 1 likewise summarizes this information. The final significant result was the interaction between interface, mode, and N-back test level (see Table 2).

For the latency during the respiratory rate detection task, the significant results were mode and the interaction between mode and N-back level. For mode ( $F = 25.584$ , d.f. = 1,  $p = 0.000007$ ), absolute mode resulted in mean latency 14.448 seconds and variance 6.56, while relative mode resulted in latency 6.949 and variance 4.19. Table 3 summarizes

**Table 1.** The latency for the heart rate task for the interactions between interface and N-back level. The first two columns contain the interface and N-back level, while the third column contains the mean latency in seconds, and the fourth column contains the variance.

Interface:	N-back:	Mean:	Variance:
Audio	N = 1	7.867	4.2
	N = 2	10.042	6.93
Vibrotactile	N = 1	12.176	6.75
	N = 2	9.125	4.03

**Table 2.** The latency for the heart rate task for the interactions between interface, mode, and N-back level. The first three columns contain the interface, mode, and N-back level, while the fourth column contains the mean latency in seconds, and the fifth column contains the variance.

Interface:	Mode:	N-back:	Mean:	Variance:
Audio	Absolute	N = 1	10	4.97
		N = 2	13.778	5.95
	Relative	N = 1	6	2.38
		N = 2	6.306	5.93
Vibrotactile	Absolute	N = 1	17.75	5.45
		N = 2	9.833	4.38
	Relative	N = 1	7.222	2.54
		N = 2	8.214	3.65

the results for the interaction between mode and N-back level ( $F = 4.195$ , d.f. = 1,  $p = 0.046$ ), which shows that absolute mode resulted in mean latency 15.865 seconds and variance 6.69 when  $N = 1$ , and latency 13.131 and variance 6.40 when  $N = 2$ . In relative mode, the mean latency was 5.600 seconds and variance 2.06 when  $N = 1$ , while the latency was 8.506 and variance 5.45 when  $N = 2$ . For the latency during the N-back test, no statistically significant results were found.

## DISCUSSION FOR STUDY TWO

### Accuracy

As with the first study, one of the largest influences on accuracy for both tasks was mode. For the heart rate detection task, the mean accuracy for relative mode was 19.84% higher, while it was 25.63% higher for the respiratory rate detection task. This result reinforces the conclusions from the first study, mainly that participants were more accurate when it was not up to them to detect the changes by monitoring each heart beat and breath. Many participants noted in both studies that it was harder to monitor respiratory rate than heart rate in absolute mode because of its less frequent events; we believe that the larger difference between the mean accuracy scores between the two modes for the respiratory task reflects this increased difficulty.

**Table 3.** The latency for the respiratory task for the interactions between mode and N-back level. The first two columns contain the mode and N-back level respectively, while the third column contains the mean latency in seconds, and the fourth column contains the variance.

Mode:	N-back:	Mean:	Variance:
Absolute	N = 1	15.865	6.69
	N = 2	13.131	6.4
Relative	N = 1	5.6	2.06
	N = 2	8.506	5.45

For N-back test accuracy, the only significant result was that participants were 7.80% more accurate when  $N = 1$ . We were surprised to find out that none of the interfaces affected this score significantly, which may indicate that participants focused more on the N-back test than the interfaces.

### Latency

Mode once again plays a significant role when considering latency for both tasks. For the heart rate task, absolute mode was 5.98 seconds longer on average, whereas the same mode was 7.50 seconds longer on average during the respiratory task. We believe that the increased difficulty of the respiratory detection task explains these data as outlined above.

For latency for the heart rate detection task, the interaction between interface and N-back test level proved to be significant. For the audio interface, the latency increased by 2.18 seconds on average when  $N$  increased from  $N = 1$  to  $N = 2$ , which intuitively makes sense since the latter test was more difficult; however, for the vibrotactile interface, the mean latency actually decreased by 3.05 seconds as the N-back test became more difficult. One plausible explanation is that as the N-back test required more attention by the participant, it became easier to separate the monitoring and N-back tasks when they used different sensing modalities (tactile and audio), rather than trying to do both tasks using only audio.

When examining the interaction between mode and N-back level for the latency data during the respiratory detection task, participant performance followed an intuitive pattern for three of the four mode-level combinations. The mode-level combination with the least amount of latency was relative mode with N-back level  $N = 1$ , followed by relative mode with N-back level  $N = 2$ . Both of those combinations had less latency than absolute mode with N-back level  $N = 1$ , which increased the latency by 7.36 seconds on average. Surprisingly, the combination of absolute mode with N-back level  $N = 2$  did not increase that latency, but rather decreased the mean latency by 2.73 seconds. We were unable to come up with a satisfying explanation for this anomaly; it is our hope that receiving additional feedback from future users of the interfaces will allow us to clarify this result.

The most puzzling result is the interaction between the interface, mode, and N-back level for the heart rate task. While we observe basic patterns, such as relative mode having lower latency than absolute mode, and there is an increase in latency as the N-back level increases, there are several pieces of data that break these logical progressions. For example, unlike previously explained results, the audio interface consistently has lower latency in relative mode compared to the vibrotactile interface in relative mode; this is at odds with the pattern observed for the two modes during the respiratory rate task. The most surprising piece of data is the extremely high latency for the vibrotactile-absolute combination when the N-back task was  $N = 1$ . The latency for that combination was on average 7.92 seconds longer than the same combination with  $N = 2$ . Since the order of the vibrotactile-absolute and vibrotactile-relative combinations was randomized when  $N = 1$ , we do not believe that it is a learning effect of the interface itself, but rather the specific vibrotactile-absolute

combination. One tentative explanation is that users have a very hard time adapting to the vibrotactile interface in absolute mode, but are quickly able to adjust by the time they experience it again; however, additional data is required to either confirm or deny this explanation.

### FUTURE WORK

Given feedback from these initial studies, we have several revisions and enhancements planned for the interfaces. We will go back and study the mapping between indicators and physiological events in more depth; for this iteration, we selected indicators that seemed intuitive to us, but may have not been optimal. The fact that two participants were unable to hear the sounds for increases and decreases in respiratory rate in relative mode for the first study indicates that we should replace the sound with something more distinctive.

After integrating the feedback that we received, we can focus on the next study. We elected to use sighted participants for these initial studies due to the practicality of finding a large enough population. Now that we have confirmation that the interfaces are viable, the next study will focus on using blind participants. An additional goal of the next study is to introduce noise into the environment when users are monitoring the interfaces to approximate everyday working conditions; this can be ambient noise for the audio interface, and additional movement for the vibrotactile interface.

### CONCLUSION

In this paper, we described the designs of two novel interfaces that allow a handler to monitor the physiological status of a canine in real-time. The application of non-visual monitoring would be useful in several fields, but we selected to use blind handlers and their guide dogs to illustrate the concept. We developed an audio and a vibrotactile interface using a BeagleBone Black as the central component to allow the monitoring of heart and respiratory rate in a discreet manner. We used a simulated canine and virtual walk system to allow experiments to be controlled, consistent, and repeatable.

We found during the first study that participants felt more comfortable with the audio-based interface, despite being just as or more accurate with the vibrotactile interface in most instances. Some participants were unable to hear some of the indicators for the audio interface in relative mode, which indicates that the interface requires revisions in order to make the sounds more distinctive. The drawback of both interfaces in relative mode is that rate of change information is not communicated to the handler, which makes estimation of heart and respiratory rate impossible. We found during the second study that relative mode consistently scores better in accuracy and latency than absolute mode. We also discovered that as an audio-based task becomes more difficult, the latency of the audio interface increases while it decreases for the vibrotactile interface. Finally, we identified some patterns between the interfaces, modes, and concentration-based task that requires additional data to understand fully.

We plan to revise the interfaces based on feedback gathered during these studies, and expand the scope of the evaluation of the interfaces. More specifically, we intend to open the

study to people with visual impairments, both with and without guide dog handling experience, to understand how prior experience may influence accuracy and latency.

## ACKNOWLEDGMENTS

The authors would like to thank Cecilia and Tammy for their valuable assistance. This material is based upon work supported by the National Science Foundation, under both Graduate Research Fellowship Grant No. DGE-1252376 and NSF grant 1329738. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

1. Gerold Baier, Thomas Hermann, and Ulrich Stephani. 2007. Multi-channel sonification of human EEG. In *Proceedings of the 13th International Conference on Auditory Display*.
2. Suzanne P Balik, Sean P Mealin, Matthias F Stallmann, and Robert D Rodman. 2013. GSK: universally accessible graph sketching. In *Proceeding of the 44th ACM technical symposium on Computer science education*. ACM, 221–226.
3. Stephen Brewster and Lorna M Brown. 2004. Tactons: structured tactile messages for non-visual information display. In *Proceedings of the fifth conference on Australasian user interface-Volume 28*. Australian Computer Society, Inc., 15–23.
4. Stephen Anthony Brewster. 1994. *Providing a structured method for integrating non-speech audio into human-computer interfaces*. Ph.D. Dissertation. Citeseer.
5. Stephen A Brewster, Peter C Wright, and Alistair DN Edwards. 1995. Guidelines for the creation of earcons. In *Adjunct Proceedings of HCI*.
6. Rita Brugarolas, James Dieffenderfer, Katherine Walker, Ashley Wagner, Barbara Sherman, David Roberts, and Alper Bozkurt. 2014. Wearable wireless biophotonic and biopotential sensors for canine health monitoring. In *SENSORS, 2014 IEEE*. IEEE, 2203–2206.
7. Rita Brugarolas, David Roberts, Barbara Sherman, and Alper Bozkurt. 2012. Posture estimation for a canine machine interface based training system. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 4489–4492.
8. Rita Brugarolas, David Roberts, Barbara Sherman, and Alper Bozkurt. 2013. Machine learning based posture estimation for a wireless canine machine interface. In *Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireless), 2013 IEEE Topical Conference on*. IEEE, 10–12.
9. Angela Chang, Sile O'Modhrain, Rob Jacob, Eric Gunther, and Hiroshi Ishii. 2002. ComTouch: design of a vibrotactile communication device. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM, 312–320.
10. Edward Childs. 2005. Auditory graphs of real-time data. In *International Conference on Auditory Display (ICAD 2005)*, Limerick, Ireland.
11. Keith M Franklin and Jonathan C Roberts. 2003. Pie chart sonification. In *Information Visualization, 2003. IV 2003. Proceedings. Seventh International Conference on*. IEEE, 4–9.
12. Myounghoon Jeon, Thomas M Gable, Benjamin K Davison, Michael A Nees, Jeff Wilson, and Bruce N Walker. 2015. Menu Navigation With In-Vehicle Technologies: Auditory Menu Cues Improve Dual Task Performance, Preference, and Workload. *International Journal of Human-Computer Interaction* 31, 1 (2015), 1–16.
13. Lynette A Jones and Nadine B Sarter. 2008. Tactile displays: Guidance for their design and application. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50, 1 (2008), 90–111.
14. KM Miller, CC Price, MS Okun, H Montijo, and D Bowers. 2009. Is the n-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology* 24, 7 (2009), 711–717.
15. Michael A Nees and Bruce N Walker. 2011. Auditory displays for in-vehicle technologies. *Reviews of human factors and ergonomics* 7, 1 (2011), 58–99.
16. Jessie YC Ng, Jo CF Man, Sidney Fels, Guy Dumont, and J Mark Ansermino. 2005. An evaluation of a vibro-tactile display prototype for physiological monitoring. *Anesthesia & Analgesia* 101, 6 (2005), 1719–1724.
17. Ivan Poupyrev, Shigeaki Maruyama, and Jun Rekimoto. 2002. Ambient touch: designing tactile interfaces for handheld devices. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*. ACM, 51–60.
18. Dongseok Ryu, Gi-Hun Yang, and Sungchul Kang. 2009. T-hive: Vibrotactile interface presenting spatial information on handle surface. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*. IEEE, 683–688.
19. Aaron E Sklar and Nadine B Sarter. 1999. Good vibrations: Tactile feedback in support of attention allocation and human-automation coordination in event-driven domains. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 41, 4 (1999), 543–552.
20. Bruce N Walker and Joshua T Cothran. 2003. Sonification Sandbox: A graphical toolkit for auditory graphs. (2003).