Hit The Road, Jack

P3: System Prototype and Evaluation Plan

Amit Garg
Rachel LeRoy
Sahib Singh
John Crisp
Bryan Bennett

Table of Contents

Description of System Prototype

The purpose of these vertical, medium-fidelity prototypes is to identify the differences between three different sensory modes of receiving messages: visual, auditory, and tactile. The prototype will consist of two separate receiving items, one for auditory displays and one for tactile displays.

The auditory displays will utilize bonephones (might use bluetooth headphones) and the tactile displays will utilize a custom vibrotactile headband. Both will be controlled by a bluetooth remote that has been programmed with simple directional cues (i.e. move forward, left, right, back, halt).

The system comprises three components: 1) the bluetooth remote control with buttons that represent directional movements, 2) a web application with a set of digital 'soundboard' style buttons that will be mapped to the buttons of the remote control, and 3) either the bluetooth headphones/bonephones or the tactile headband, which will communicate with the web app and initiate a corresponding, pre-programmed message to the participant either through auditory or tactile displays.

Description of User Interaction

Both prototypes will be device-initiated interactions because we are focusing on how participants passively receive information via our prototypes.

Auditory Display

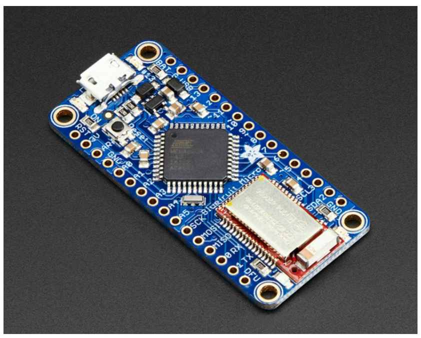
Users will receive auditory displays by wearing bonephones. Bonephones are like mini speaker drivers worn on the temporal bones. Away from the head, they produce inaudible sound, but when placed on the temporal bones they use the space of the human head to amplify the sound. Sound travels directly to the inner ear to be perceived. We chose this method because it allows for the ear canal to remain open and for the participant to remain situationally aware. However, earbud headphones might be used out of convenience since we are interested in testing different sensory modes, not different types of auditory communication. If auditory displays prove more efficient than tactile displays, then a true bonephone will be acquired and prototyped with a higher fidelity. The bonephone will be an output method, delivering the message sent by the remote to the receiving participant. Other than turning the bonephone on, wearing the bonephone, listening to the message, and performing some directional movement related to the message, no other interaction will occur with the prototype.

Tactile Display

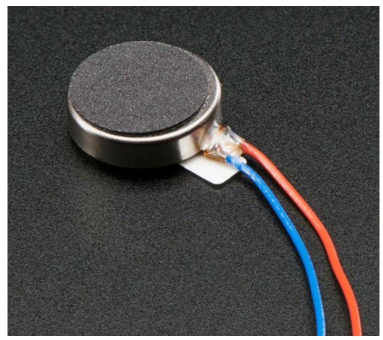
Users will receive tactile displays by wearing a custom vibrotactile headband around their head. Currently, no commercial product can be adapted for this purpose. Therefore, the vibrotactile headband needs to be built by our team. Our design uses 6 Adafruit vibrating mini motor discs that are sewn onto a headband and controlled by an Adafruit Bluefruit LE

Micro which has a powerful microcontroller and a built-in bluetooth module for wireless communication with a remote control. For example, to signal “move forward”, the “forward” button on the remote would be pressed. Then, a preprogrammed vibrotactile pattern would initiate with vibrations at the back of the headband and move along the discs toward the front of the headband. Other than turning on the headband, wearing the headband, feeling the message, and performing some directional movement related to the message, no other interaction will occur with the vibrotactile headband.

Adafruit Bluefruit LE Micro



Vibrating Mini Motors



Implementation Challenges

The biggest challenge we faced was realizing that our intended subject population - the ROTC cadets - are a protected population by the Department of Defense (DoD). With a DoD population, more stringent and thorough protocol review is required before testing can commence. In light of this discovery, we have pivoted the project to encompass a different subject population so that our device can begin to be tested while we wait for DoD approval. Therefore, we submitted an additional IRB protocol that proposes the testing of a team-building communication game to be tested by Georgia Tech students. The game would impose the low-light, no talking constraints that the ROTC population works in.

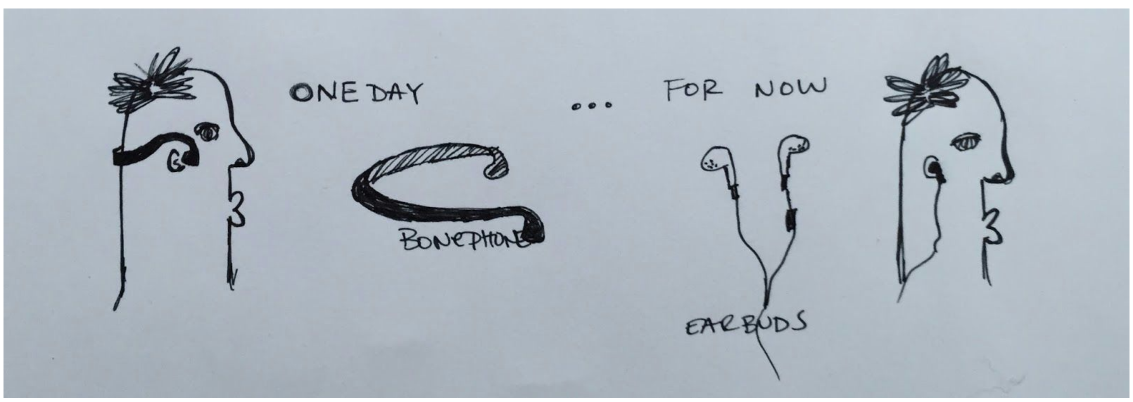
Another difficulty we encountered was how to set up a viable evaluation plan to test metrics that related to our original objective - increasing efficiency of ground movement in low-light situations. Since the communication model involves both sending and receiving devices, we were unsure of how to test both in a way that could help us narrow down which devices would be more intuitive. Every evaluation plan we came up with led us to obvious answers or was not effective in eliminating biases associated with already learned activities. For example, testing hand and arm signals against a new input method would undoubtedly show a faster send time for hand and arm signals because these are already learned by our initial subject population. However, this realization was actually useful because it helped us recognize that we should not reinvent the wheel. If hand and arm signals work in the day time, why change them at night? The reason hand and arm signals are ineffective at night is because they rely on line of sight as the means of perception and understanding. So if the means of perceiving a message at night can be changed to tactile or auditory feedback, then hand and arm signals can remain relevant in both high and low light scenarios. This eliminated the need to prototype and test novel sending devices and helped us focus only on the receiving side.

The evaluation plan has led the decision-making as it relates to receiving device designs. In order to evaluate the different methods (visual, auditory, and tactile), mid-fidelity prototypes that can quickly be prototyped and revised were developed. Once quantifiable data has been gathered on each reception method, we can then make a decision on which reception method to implement in the final design.

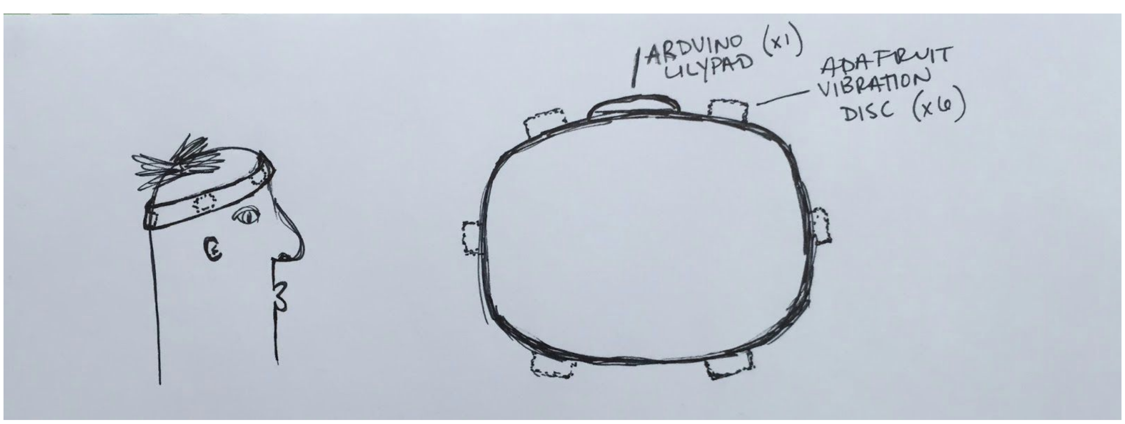
Note: the bluetooth remote was chosen as a means of initiating the auditory and tactile displays because a team member had already been working on such a device for a separate research project at Georgia Tech. It utilizes an Adafruit Bluefruit EZ-key for easy mapping of button outputs.

Photos & Sketches:

Auditory Display - Bonephone / Headphone

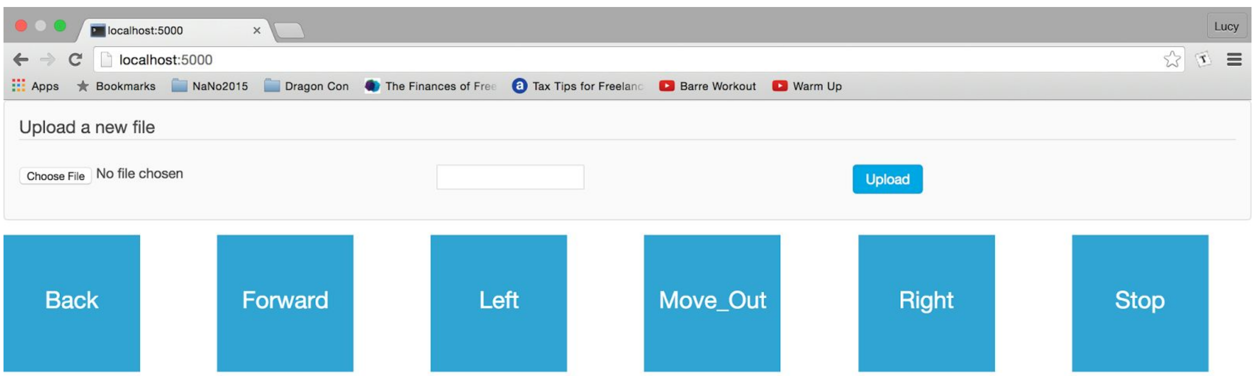


Tactile Display - Vibrotactile Headband



Input Device - 3D Rendering of Bluetooth Remote

Web App Audio Interface - Function To Be Connected to The Remote and Prototype



Usability Specifications for Initial Prototype

The purpose of these vertical, medium-fidelity prototypes is to identify the differences between three different sensory modes of receiving messages: visual, auditory, and tactile.

The following will be consistent during testing:

- the participant will be trained in the use of the prototype at the beginning of the assessment
- sending will be controlled by a member of our team, who will operate the prototyped sending device (i.e. the remote control in auditory and tactile groups, hand/arm signals in visual group).
- the prototype assessments will be conducted in low light
- quantitative information will be gathered during the assessment, measuring the prototypes' **detectability, learnability** and **efficiency** in providing information to the receiving participant during a task
- qualitative information will be gathered at the end of the assessment, measuring the **usability** of the prototypes as well as the participants' perceived **task load** while

interacting with the prototypes. This will be done with modified versions of the System Usability Scale (SUS) and Task Load Index (TLX) by NASA (see *Appendix A*).

The appropriate usability criteria (tasks a user should be able to complete, and how they will be evaluated) are further defined below for each of the three research groups..

Auditory tasks:

- the participant should accurately detect that an auditory instruction was transmitted - **detectability**
- the participant should respond accurately to auditory instructions - **learnability**
- the participant should respond in a timely manner to auditory instructions - **efficiency**

Auditory evaluation:

- detectability will be measured based on whether or not the participant reacts to the display.
- learnability will be measured by the number of errors that occur during the task
- efficiency will be measured by the time it takes the participant to finish the entire task, as well the time between the leader sending a message and the participant completing the correct corresponding movement (i.e. response time, RT).
- participants will take surveys at the end of the assessment to measure perceived usability and task load during interaction with the prototype.

Tactile tasks:

- the participant should accurately detect that a tactile instruction was transmitted - **detectability**
- the participant should respond accurately to tactile instructions - **learnability**
- the participant should respond in a timely manner to tactile feedback - **efficiency**

Tactile evaluation:

- detectability will be measured based on whether or not the participant reacts to the display
- learnability will be measured by the number of errors that occur during the task
- efficiency will be measured by the time it takes the participant to finish the entire task, as well the time between the leader sending a message and the participant completing the correct corresponding movement (RT).
- participants will take surveys at the end of the assessment to measure perceived usability and task load during interaction with the prototype.

Visual tasks:

- the receiving participant should accurately detect that a visual signal was transmitted - **detectability**
- the receiving participant should respond accurately to a visual signal - **learnability**
- the receiving participant should respond in a timely manner to a visual signal - **efficiency**

Visual evaluation:

- detectability will be measured based on whether or not the participant reacts to the
- learnability will be measured by the number of errors that occur during the task
- efficiency will be measured by the time it takes the participant to finish the entire task, as well the time it takes between the leader sending a message and the participant completing the correct corresponding movement (RT).
- participants will take surveys at the end of the assessment to measure perceived usability and task load during interaction with the prototype.

A potential evaluation plan for a full, send-and-receive system is described in *Appendix B: Usability Specifications for a Fully Envisioned System.*

Initial Evaluation Plan

Our user testing session will involve testing different sensory modes of receiving information within a communication paradigm. We are going to test the efficiency differences between 3 different sensory modes of receiving messages: visual, auditory, and tactile. There will be 3 research groups:

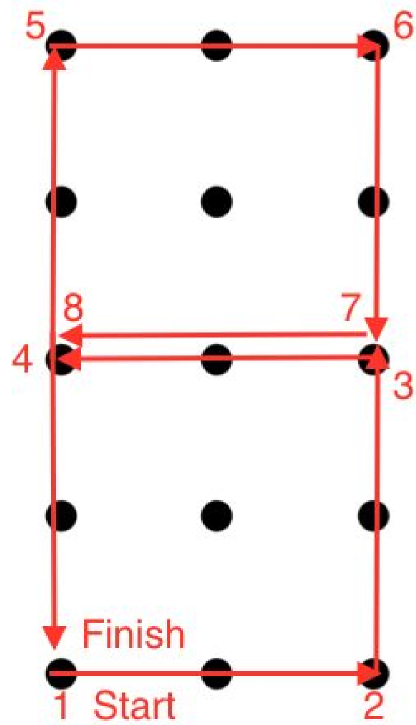
1. visual, receiving hand and arm signals via line of sight (LOS).
2. auditory, receiving messages via bone/headphones.
3. tactile, receiving messages via a tactile headband.

In the evaluation plan, each research group will be the same except for the method of receiving information (e.g. visual, auditory, tactile). All other factors will be constant. Therefore, our defined independent variable is sensory mode for receiving information, while our dependent variables are total task time, response time (RT), number of errors, perceived usability, and perceived task load. We hypothesize that there will be a difference between the three sensory modes for receiving information along all of our dependent variables.

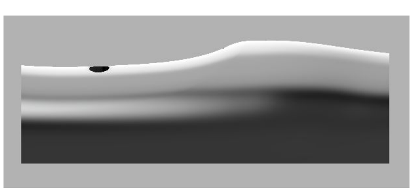
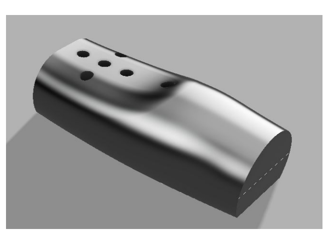
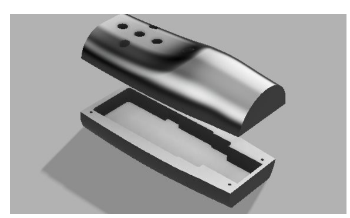
The layout of the evaluation system will be as follows:

- Each group will have a leader (from our team) and a follower (participant). The leader and follower will engage in a very simple obstacle course in the dark, hopefully outdoors.
- The obstacle course will have a set of 15 soccer cones, in grid format.

- The leader/follower will walk through the grid systematically through the shape of a digital 8. Only the leader will know when to send commands to the follower, and will have a predefined set of commands for each trial.
- We chose 15 cones so that the follower is not able to guess what is happening while on the course, and a digital 8 because we can give a variety of directional cues with that shape.
- We will collect the time it takes to complete the entire obstacle course, as well as the time it takes to complete each action within the obstacle course (e.g. RT)
- We will also administer post-testing surveys to assess the usability, using the System Usability Scale (SUS), and the perceived task load, using NASA's Task Load Index (TLX). We will modify these surveys according to our situation.



The bonephones and tactile headband will both be controlled by a small, handheld remote control. The leader will press a button on the remote, it will communicate to a device via bluetooth, and the device will send pre-recorded audio messages to the bonephones or pre-programmed tactile patterns to the tactile headband, depending on the research group.

Appendix A: Task Load Index (TLX) and System Usability Scale (SUS)

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand How mentally demanding was the task?

Very Low Very High

Physical Demand How physically demanding was the task?

Very Low Very High

Temporal Demand How hurried or rushed was the pace of the task?

Very Low Very High

Performance How successful were you in accomplishing what you were asked to do?

Perfect Failure

Effort How hard did you have to work to accomplish your level of performance?

Very Low Very High

Frustration How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low Very High

	Strongly disagree				Strongly agree	
	1	2	3	4	5	
1. I think that I would like to use this system frquently.					√	4
2. I found the system unnecessarily complex.				√		1
3. I thought the system was easy to use.		√				1
4. I think I would need the support of a technical person to be able to use this system.	√					4
5. I found the various functions in this system were well integrated.		√				1
6. I thought this system was too inconsistent.			√			2
7. I would imagine that most people would learn to use this system very quickly.		√				1
8. I found the system very cumbersome to use.				√		1
9. I felt very confident using the system.					√	4
10. I needed to learn a lot of things before I could get going with this system.		√				3

Appendix B: Usability Specifications for a Fully Envisioned System

The fully envisioned system will include sending and receiving solutions that are connected via a transmission medium. The usability specifications for the eventual sending solution should be practically identical to the usability specifications for the receiving prototype as described above. This section will focus on describing usability specifications for the sending solution of the fully envisioned system.

The following will be consistent when assessing the sending solution:

- the sending participant will be trained in the use of the sending solution at the beginning of the assessment, and will operate the sending device
- the usability assessment will be conducted in low light
- quantitative information will be gathered during the assessment, measuring the sending device's effectiveness and efficiency in sending instructions
- qualitative information will be gathered at the end of the assessment, measuring the sending participant's satisfaction with the usability of the sending device.

The appropriate usability criteria (tasks a user should be able to complete, and how they will be evaluated) are identified below.

Sending tasks:

- the sending participant should select a message in a timely manner in response to a stimulus
- the sending participant should accurately select the intended message in response to a stimulus
- the sending device should be able to accurately transmit a message in a timely manner

Sending evaluation:

- we would measure the time it takes the sending participant to select a message once a stimulus is presented
- we will assess the accuracy with which the sending participant selects the correct message in response to a stimulus
- the sending participant(s) will be interviewed at the end of the assessment to measure usability and task load. Assessments will include measures of detectability, learnability, and efficiency.
- the sending device's battery life should be sufficient, measured as a percent of expected use. The minimum expected use would be for an overnight mission (10-12 hours), however, possibility of extended time low visibility scenarios (Arctic Circle in winter; jungle; smog/smoke/fire) should be considered.
- the sending device's weight should be judged appropriate, as measured by a survey of the sending participant having worn the device for at least 30 minutes

- the sending device's apparent ruggedness should be appropriate, as subjectively measured by a survey of the sending participants