Exploring Feedback Modalities Using Wearable Device for Complex Systems Training Programs

Layla Akilan
Wright State University

Follow this and additional works at: https://corescholar.libraries.wright.edu/etd_all

Part of the Operations Research, Systems Engineering and Industrial Engineering Commons

https://corescholar.libraries.wright.edu/etd_all/2237

This Thesis is brought to you for free and open access by the Theses and Dissertations at CORE Scholar. It has been accepted for inclusion in Browse all Theses and Dissertations by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.
EXPLORING FEEDBACK MODALITIES USING WEARABLE DEVICE FOR COMPLEX SYSTEMS TRAINING PROGRAMS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Industrial and Human Factors Engineering

By

Layla Akilan
B.S., Wright State University, 2015

2018
Wright State University
WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

DECEMBER 14, 2018


_________________________________
Subhashini Ganapathy, Ph.D.
Thesis Director

_________________________________
John Gallagher, Ph.D.
Chair, Department of Biomedical, Industrial and Human Factors Engineering

Committee on Final Examination

_________________________________
Subhashini Ganapathy, Ph.D.

_________________________________
Mary E. Fendley, Ph.D.

_________________________________
Sasanka V. Prabhala, Ph.D.

_________________________________
Barry Milligan, Ph.D.
Interim Dean of the Graduate School
ABSTRACT


This study examined the effectiveness of a wearable device in delivering various feedback modalities in an attempt to improve performance outcomes in complex systems. Secondarily this study looked at performance when feedback type was matched to preferred learning style according to VARK Learning Styles Inventory results.

Participants were required to perform system monitoring and correct for system failures through key presses. Feedback was delivered through a smart watch and was based on response time performance. Feedback modalities included visual, auditory, and haptic feedback. Subjective ratings of situation awareness and mental workload were also examined. Results indicated that auditory feedback condition response times were significantly slower than response times in other feedback condition with the control group having the fastest mean response times. Participants who tested as read write learners were the only learning style group to show higher levels of situation awareness and decreased mental workload when presented with their preferred instructional style.
# TABLE OF CONTENTS

1. **BACKGROUND** ................................................................................................................................. 1  
   1.1 Understanding Complex Systems in the Military Domain .................................................. 1  
   1.2 Dynamic Systems ............................................................................................................................ 2  
   1.3 Situation Awareness, and Mental Workload ............................................................................. 3  
   1.4 Feedback .......................................................................................................................................... 3  
   1.5 Wearable Devices ............................................................................................................................. 4  
   1.6 Learning Style ................................................................................................................................. 4  
   1.7 VARK Learning Style Inventory ................................................................................................. 5  

2. **RESEARCH OBJECTIVES** ............................................................................................................... 6  

3. **METHODOLOGY** .............................................................................................................................. 7  
   3.1 Participants ...................................................................................................................................... 7  
   3.2 Materials ......................................................................................................................................... 7  
   3.3 Design ............................................................................................................................................ 8  
   3.4 Experimental Task ........................................................................................................................ 9  
   3.5 Procedure .................................................................................................................................... 13  

4. **RESULTS** ........................................................................................................................................ 14  
   4.1 VARK Scores ................................................................................................................................. 14  
   4.2 Effect of Feedback Modality on Response Time ....................................................................... 15  
   4.3 Feedback Notifications per Condition ....................................................................................... 16  
   4.4 Feedback Notifications Over Time ............................................................................................. 17  
   4.5 Situation Awareness Relative to Feedback Modality ............................................................... 19  
   4.6 Situation Awareness when Learning Style Matched Instructional Style ............................. 20  
   4.7 Mental Workload Scores Relative to Feedback Modality ........................................................ 21  
   4.8 Mental Workload when Learning Style Matched Instructional Style .................................. 22  
   4.9 System Usability ........................................................................................................................... 23  

5. **DISCUSSION** .................................................................................................................................. 23  
   5.1 Response Times ............................................................................................................................. 23  
   5.2 VARK ............................................................................................................................................ 24
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Desired states for system monitoring F1-F8</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Light failure in F6</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Gauges failure in F1</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Pump failure in F7</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>Watch Interface Design</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>Experimental Set Up</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>VARK Score Distribution</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>Visual Feedback Trials: Number of Feedback Notifications per System Failure.</td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>Haptic Feedback Trials: Number of Feedback Notifications per System Failure.</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>Auditory Feedback Trials: Number of Feedback Notifications per System Failure.</td>
<td>18</td>
</tr>
<tr>
<td>11</td>
<td>Mixed Modal Feedback Trials: Number of Feedback Notifications per System Failure.</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>Mean SA scores per feedback condition.</td>
<td>21</td>
</tr>
<tr>
<td>13</td>
<td>Mental workload for each Feedback Condition</td>
<td>22</td>
</tr>
</tbody>
</table>
1. BACKGROUND

Wearable devices have the potential to improve performance, increase situation awareness, and decrease mental workload, (Ziegler et al., 2015). In particular, haptic feedback has shown promise in being an effective sensory modality to deliver feedback, (Prewett et al., 2006). Wearable devices may be a promising set of tools which can be used to enhance training programs in military domains. In this study we examine different modes of feedback delivered through a smart watch, how they affect performance, situation awareness, and mental workload in complex systems, and secondarily we examined the relationship between feedback and preferred learning style relative to response time performance.

1.1 Understanding Complex Systems in the Military Domain

It has been a longstanding effort by the United States military to support and fund efforts related to increasing intelligence. According to Noble, (1989), this continued support and funding can be mapped to practical reasons which include, developing intelligent and autonomous systems, accelerating instructional efforts and learning outcomes in training, and amplifying interactions between humans and machines. Exploring ways in which we can improve training programs in complex systems, has the potential to improve performance and safety overall for military personnel. These efforts are also necessary in order to keep up with a quickly evolving technological landscape. Leveraging mobile devices and the affordances they offer for delivering feedback effectively and efficiently, may be one way to improve training program outcomes. The following sections will
outline the meaning of dynamic and complex systems, the importance of situation awareness and mental workload in terms of performance, and the potential for wearable devices to deliver meaningful feedback that improve training outcomes. Additionally, a secondary focus of this study, considering individual differences in learners preferred learning style, is described in relation to the VARK Learning Style Inventory.

1.2 Dynamic Systems

The nature of dynamic and or complex systems consists of making many decisions within a small amount of time, requires understanding the state of an environment that is constantly changing, requires constant and continuous analyses to understand the environment, and requires the ability to interpret the meaning of the environment relative to pertinent goals, (Endsley, 1995). Pilots, for example, are expected to understand and follow standard operating procedures, some of which include; understanding the airspace, understanding the routes of other aircraft, accounting for sudden changes in airspace and flight plan, remaining cognizant of weather conditions and fuel levels, and having the ability to recognize mechanical failures, (Lin & Lu, 2016). All of this is done, while simultaneously adjusting their understanding of the environment to their goals. Mental workload can be high in these scenarios and maintaining situation awareness becomes exponentially more difficult as the complexity and dynamics of the environment increases. Most accidents in aviation are the result of a lack of Situation Awareness (SA), (Lin & Lu 2016).
1.3 Situation Awareness, and Mental Workload

Situation Awareness refers to how one understands their environment and higher levels of SA have been correlated with improved performance, (Ikuma, Harvey, Taylor & Handal, 2014). According to Endsley’s “model of situational awareness in dynamic decision making,” the output of SA is decision-making, (1995, p.35).

Mental Workload (MWL) pertains to the complexity of a task and the mental demand associated with that task. When the demand of any given task exceeds available resources related to attention, this can result in reduced effectiveness, decision-making, and performance, (Lin & Lu 2016). To improve training programs, learning outcomes, and long-term performance of operators, it is necessary to design meaningful human machine interactions and feedback for training scenarios, that has the ability to improve an operator’s understanding of the world without simultaneously adding complexity and increasing mental demand.

1.4 Feedback

It has been reported that outcomes related to learning are highly moderated by characteristics specific to individual learners such as motivation and that feedback has been demonstrated as a strong motivator (Shute, 2008). Feedback assists in scaffolding, which allows us to engage in higher-level thinking and problem solving where it would not be possible without such assistance due to lack of experience and knowledge, (Shute, 2008). Immediate and even delayed forms of feedback have shown to improve performance on learning tasks, (Dobryakova & Tricomi, 2013).
Control theory tells us that every system has a desired state, a medium in which we can control the state of the system, and a feedback loop, which allows us to compare the current state to the desired state, (Carver & Shier, 1990). This control loop involves a human’s ability to use feedback to correct errors and make changes to the system based on its desired state, (Norman, 1990). Feedback is an important tool that helps us adjust our goals and strategies during learning, (Arbel & Wu, 2016).

1.5 Wearable Devices

There is evidence in the literature that suggests that wearable devices that utilize haptic feedback can improve SA and performance (Wolf & Kuber, 2017). Haptic feedback, has been found to have a significant effect on human effectiveness, (Prewett et al., 2006). Smart watches in particular have been shown to improve SA and reduce MWL, through the use of various kinds of feedback (Ziegler et al., 2015). We believe that feedback modalities in wearable devices should be explored in an attempt to better understanding the kind of tools and devices that can be leveraged in order to improve training programs for dynamic systems.

1.6 Learning Style

For the purpose of this study, learning style refers to the individual differences among learners pertaining to their preferred mode of instruction, (Pashler, 2009). Meshing theory states that it is optimal to pair learning style with instructional style because it plays on the strengths of learners in terms of cognitive ability (Pashler, 2009). The VARK learning styles inventory, invented by Fleming, closely aligns with Meshing Theory, in that it is focused on revealing a learner’s preferred method of gathering information during
learning through sensory input (Fleming, 1995). VARK stands for Visual, Aural, Read/write, and Kinesthetic and represents the different types of learners that result from taking the survey. Visual learners are said to prefer symbolic visual representations, aural learners are said to prefer auditory modes of instructions, read write learners are reported as preferring written modes of visual input, while kinesthetic learners prefer mixed modes of sensory input. A study performed by Leite, Svinicki, & Shi established validity for the VARK and its ability to predict learning style with a sample size of greater than 15,000 participants, in 2009.

1.7 VARK Learning Style Inventory

The VARK Learning Styles Inventory was chosen for this study. Read write learners could be provided with written text as visual feedback, aural learners with auditory feedback, and kinesthetic learners with mixed modal feedback including haptics. This mapping allowed for us to evaluate whether or not individual differences, such as preferred learning style, had any effect on performance in the different feedback conditions.

Overall, we believe that there may be opportunities to improve training programs through the use of wearable devices such as smart watches. The technology surrounding smart watches continues to improve and the future looks promising for smart watch capabilities. We investigated the different modalities offered by smart watches and how they affect response time performance in a complex system test bed. Secondarily we attempted to examine individual’s preferred sensory input in relation to feedback modalities offered by the watch.
2. RESEARCH OBJECTIVES

The following research objectives were defined in order to draw conclusions about the different feedback modalities relative to response time performance, situation awareness, mental workload, and learning styles.

1. How do different feedback modalities, delivered through a wearable device, affect performance when measured as response time to system failures?

   \( H_0: \) There is no significant difference between feedback modalities as it pertains to response time.

   \( H_1: \) There will be a significant difference between feedback modalities as it pertains to response time.

2. How are subjective ratings of situation awareness and mental workload affected by different modalities of feedback?

   \( H_0: \) There is no significant difference between feedback modalities as it pertains to subjective ratings of situation awareness.

   \( H_1: \) There will be a significant difference between feedback modalities as it pertains to subjective ratings of situation awareness.

3. How does learning style affect performance when learners receive their preferred type of feedback?

   \( H_0: \) There is no significant difference between learners when they receive their preferred feedback type, as it pertains to response time.
3. METHODOLOGY
The following methodology was utilized in an attempt to create a complex system test bed where performance could be measured as response time to system failures and feedback modalities could be delivered through a smart watch in real-time based on performance.

3.1 Participants
Wright State University’s Institutional Review Board approved this study. Participants were recruited from Wright State University, N=34. Participants were recruited by word of mouth and email. All students were compensated 50 USD for completing the study.

3.2 Materials
Consent forms were obtained by all participants prior to the start of each experiment followed by VARK Learning Styles Inventory to test individual learning styles. A simulated version of the Multi Attribute Task Battery (MAT-B) was built for this study in Visual Studio. MAT-B is a complex system that is designed to be consistent with the kind of tasks aircraft crews complete as part of their standard operating procedures or SOP’s. Tasks include tracking, system monitoring, resource management, and call signs and are monitored simultaneously by users. MAT-B is designed to evaluate operator performance and mental workload, (MAT-B II, 2018). It is important to note that this system is commonly used on non-pilot populations to test mental workload according to the MAT-
B official website, (MAT-B II, 2018). This system was ideal for our study. However, due to constraints in the MAT-B system, a mocked-up version had to be created in Visual Studio in order to meet the needs of our study which required integrations with the LG smart watch on the Android platform and Active MQ software so that feedback notifications could be sent in real time based on responses to system failures. Users interacted with MAT-B on a desktop computer using keyboard buttons F1, F2, F3, F4, F5, F6, F7, and F8 to correct system failures. Participants were required to sit at a desk and interact with the interface while wearing a smart watch, the Samsung LG Sport 2, in all trials. Active MQ, an open source messaging software was used to deliver feedback through the watch in real time based on these performance measures related to correcting system failures. Situation Awareness Rating Technique (SART) was used to measure subjective ratings of SA after each trial. NASA-rTLX was used to measure subjective ratings of MWL after each trial. System Usability Scale (SUS) was used to measure the overall usability of the system and was only administered at completion of the study. A second monitor adjacent to the test bed, was set up for experimenters to manipulate conditions between trials in order to carry out the random assignment.

3.3 Design

This experiment was a within subjects design. Dependent variables included task performance, specifically response time to system failures, situation awareness and mental workload scores. The primary independent variable was feedback type which included, visual, visual with auditory, visual with haptic, mixed which included visual, auditory, and haptic feedback, or control which received no feedback through the watch. Learning style was a secondary independent variable that was examined. Learning style
included 4 levels based on results of the VARK: visual, auditory, read/write, and kinesthetic. To avoid learning effects, the order in which system failures were presented in each scenario was randomized per participant using the traditional Latin Squares method. In addition, the order of the 5 feedback conditions were randomized for each participant.

3.4 Experimental Task

Participants were required to monitor a simulated version of the MAT-B user interface. The simulated version included 3 areas for system monitoring including: Gauges, lights, and pumps. During training, users were taught how to identify the desired state for each system, how to identify a system failure, and how to correct system failures through pressing the F1-F8 keys on the keyboard. Desired states can be seen below in Figure 1. where lights are green to represent an on state, gauges which can be seen below lights have tickers centered at 0, and pumps 7 and 8 on the right side have green liquid centered between 2 red threshold lines. Examples of failure states for all 3 systems can be seen in Figures 2., 3., and 4. below.
Figure 1. Desired states for system monitoring F1-F8

Figure 2. Light failure in F6

Figure 3. Gauges failure in F1
System failures occurred every 10-12 seconds. Trials contained 9 system failures in total over a 2-minute period. Feedback was designed to be delivered based on performance. Feedback notifications were sent to the watch exactly 2 seconds after the onset of a system failure, if no response from the participant was recorded. If participants responded to a system failure before the 2 second mark, feedback would not be sent to the watch. Visual feedback from the watch was presented in the watch interface as a white background with black text for maximum contrast, as seen in Figure 5 below.
Notifications were designed to display F1, F2, F3, F4, F5, F6, F7, or F8, and this text corresponded to the location of the system failure within the interface. Location information was provided as F key indicators on the watch interface in order to reduce interactions with the watch, and potentially keep the time it would take to interpret the meaning of notifications to a minimum. Only one system failure was triggered at any given time and only one notification was displayed on the watch at any given time. All conditions with the exception of the control condition were paired with this visual indicator of location for consistency. Auditory feedback consisted of the visual indicator paired with a single tone. Due to constraints with the smart watch, the tone was played through a blue tooth speaker directly in front of the participant and timing of the tone was synced with delivery of the visual message. Haptic feedback was paired with the visual indicator and consisted of a single vibration through the watch to the wrist. Participants were required to have one hand on the keyboard and to keep the hand wearing the watch free and in their peripheral vision. An example of the experimental set up can be seen in Figure 6 below.

Figure 6. Experimental Set Up
3.5 Procedure

After consent forms were obtained, participants were administered the VARK learning styles inventory to test their learning style. Learning Style scores included a distribution of scores among 4 categories, which included scores for Visual, Auditory, Read/Write, and Kinesthetic categories. Participants were then instructed to put on a smart watch, the Samsung LG Sport 2, which they would wear for the duration of the experiment.

Participants completed a 2-minute trial for each of the 5 feedback conditions. Conditions included visual feedback, visual with auditory feedback, visual with haptic feedback, visual, auditory, and haptic feedback as a mixed modal condition, and a control condition that received no feedback from the watch. Training was approximately 5 minutes long.

During training participants were exposed to what the desired state and failure state looked like for gauges, lights, and pumps. All three areas had to be monitored simultaneously on the screen. All system components contained labels for which F key would fix the failure and these were labeled directly on the user interface. If participants fixed a system failure within 0-2 seconds, this was counted as a correction and did not trigger feedback to be sent to the watch. If no response was received feedback was sent to the watch at the 2-second mark. If the participant did not respond within 12 seconds after the onset of the system failure, this was counted as a miss and the failure would time out and return to its desired state. Only one system failure was triggered at a time during any given trial. Latin Squares methods were used to randomize the order of conditions as well as the order of system failures within the 3 systems in each condition. Participants completed 5 two-minute trials. At the end of each 2-minute trial, participants completed SART and NASA-tlx questionnaires. At the end of the experiment, all participants
completed a SUS questionnaire for a total of 12 questionnaires per participant. The testing duration was approximately 1 hour. Participants were compensated in the amount of 50 USD.

4. RESULTS

4.1 VARK Scores

VARK scores were obtained by totaling the score for each category based on answers that were encoded as either; Visual, Auditory, Read/Write, or Kinesthetic. Scores were distributed across the 4 categories. The highest of the 4 scores was used to classify each participant as Visual, Auditory, or Kinesthetic learners.

The distribution of VARK scores across the 4 learning style types can be seen below in Figure 7.

![VARK Score Distribution](image)

Figure 7. VARK Score Distribution
4.2 Effect of Feedback Modality on Response Time

An ANOVA was performed (N=34) to examine the relationship between response time and the 4 levels of the independent variable feedback, which included visual feedback, visual paired with auditory feedback, visual paired with haptic feedback, and mixed modal feedback which included visual, auditory and haptic feedback, as well as a controlled condition which received no feedback. Response times received in less than 2 seconds, did not receive feedback notifications, and these data points were removed from our sample. An analysis of variance showed that the effect of feedback modality on response time was significant, $F(4,154) = 7.67, p = .0001$. Post hoc analyses using Tukey's indicated that the average response time was significantly slower in the auditory condition ($M = 4.58, SD = 0.26$) than in the other feedback conditions including haptic ($M= 3.28, SD=0.2$), mixed modal ($M=3.24, SD=0.24$), visual ($M = 3.20, SD = 0.25$), and control ($M= 2.66, SD=0.25$)

An ANOVA was performed for each group of learners to examine whether or not response time was affected when preferred learning style matched the feedback modality. Results of the VARK Learning Style Inventory (N=34) were as follows: visual learners (n-1= 5), auditory learners (n-1=5), read/write learners (n= 5), kinesthetic learners (n-5 = 5) and mixed learners, (n-2=5). An analysis of variance showed that the effect of being a visual learner was not significant relative to response times in any feedback condition, $F(4,20) = 1.23, p = 0.33$. An analysis of variance showed that the effect of being an auditory learner was not significant relative to response times in any feedback condition, $F(4,20) = 1.69, p = 0.19$. An analysis of variance showed that the effect of being a read write learner was not significant relative to response times in any feedback condition, $F
(4,20) = 1.82, p = 0.17. An analysis of variance showed that being a kinesthetic learner had a significant effect on response times between feedback conditions, F (4,20) = 5.11, p = 0.0053. Post hoc analyses using Tukeys indicated that for kinesthetic learner’s, response times in the auditory condition (M = 5.47, SD = 0.58) were not significantly different than the haptic condition (M = 3.38, SD = 0.58) but were significantly different than visual (M = 2.70, SD = 0.58), mixed (M = 2.49, SD = 0.58), and control conditions (M = 2.19, SD = 0.58).

**4.3 Feedback Notifications per Condition**

The auditory fb condition received (M = 2.72, SD=.42) feedback notifications per trial, the visual feedback condition received (M = 5.84, SD=.40) feedback notifications per trial, the control group received (M =5.97, SD=.40) feedback notifications per trial, the mixed modal feedback condition received (M = 6.09 SD= .39), and the haptic feedback condition received an average of (M = 6.56, SD= .40) feedback notifications per trial. An ANOVA was performed (N=34) to examine the relationship between number of feedback notifications per trial and the feedback condition including visual, visual with auditory, visual with haptic, mixed modal including visual auditory and haptic feedback, and a control condition which received no feedback. An analysis of variance showed that the effect of feedback modality on number of feedback notifications per trial was significant, F (4,154) = 13.62, p = .0001. Post hoc analyses indicated that the auditory feedback condition had significantly less notifications than the other feedback conditions with (M = 2.72), mixed modal being (M=6.09), visual being (M = 5.84), control being (M= 5.97) and where haptic was (M= 6.56).
4.4 Feedback Notifications Over Time

Finally, the distribution of feedback notifications seen over the course of the 9 system failures in each scenario were examined for each feedback condition. This included visual, visual+ auditory, visual+ haptic, and mixed modal feedback. While the order of system failures was randomized, all scenarios delivered system failures 1-9 every 10-12 seconds. Figures 8, 9, 10, and 11 below show the distribution of system errors over the course of the 2-minute trial for all participants in the Visual, Haptic, Auditory, and Mixed modal conditions, respectively.

Figure 8. Visual Feedback Trials: Number of Feedback Notifications per System Failure.
Figure 9. Haptic Feedback Trials: Number of Feedback Notifications per System Failure.

Figure 10. Auditory Feedback Trials: Number of Feedback Notifications per System Failure.
4.5 Situation Awareness Relative to Feedback Modality

SART scores were obtained by first summing questions pertaining to understanding, then summing questions pertaining to demand, and finally summing questions pertaining to supply. Final SART scores were obtained using the following formula: $SA = U - (D - S)$. Mean SART scores for each feedback condition were compared and were examined in groups where preferred learning style matched feedback type. Results showed the mean SA score for the visual with auditory feedback condition (M=24.38), visual with haptic feedback condition (M=24.08), visual feedback condition (M=23.44), control condition (M=23.33), and mixed modal feedback condition (M=22.15).
4.6 Situation Awareness when Learning Style Matched Instructional Style

The mean SA score was slightly higher for read write learners in the visual feedback condition with (M=22.6), where visual learners were (M=22.5), kinesthetic learners were (M=22), and auditory learners were (M=19.2) in the visual condition. The mean SA score for aural learners in the auditory feedback condition was (M=22), where read write learners were (M=23.6), visual learners were (M=22.7), and kinesthetic learners were (M=20.1) in the auditory feedback condition. The mean SA score for kinesthetic learners in the mixed modal feedback condition was (M=18.3), read write learners were (M=23), visual learners were (M=18.83), and aural learners were (M=23) in the mixed modal feedback condition. Finally, learners who tested as a two-letter combination VARK score, also known as a mixed learner, was examined relative to the mixed modal feedback condition, as all participants who tested as mixed learners happened to be some variation combined with the kinesthetic learning style. The mean SA score for mixed learners in the mixed modal feedback condition was (M=25.92). Mean SART scores for each feedback condition can be seen below in Figure 12.
4.7 Mental Workload Scores Relative to Feedback Modality

NASA-TLX raw scores were obtained by totaling the rating for each of the 6 questions. Scores were out of 100 for each question for a total of 600 possible points. Mean NASA-rTLX scores for each feedback condition were compared and were examined in groups where preferred learning style matched feedback type.

Results showed the mean NASA-rTLX scores for the visual paired with auditory feedback condition was (M=128.88), visual with haptic feedback condition (M=130.89), visual feedback condition (M=113.87), control condition (M=119.48), and mixed modal feedback condition (M=154.28). Mean NASA-rTLX scores for each feedback condition can be seen below in Figure 13.
4.8 Mental Workload when Learning Style Matched Instructional Style

The mean NASA-rTLX scores for read write learners in the visual feedback condition with \((M=85.4)\), where visual learners were \((M=138)\), kinesthetic learners were \((M=84.3)\), and auditory learners were \((M=180.8)\) in the visual condition. The mean NASA-rTLX scores for aural learners in the auditory feedback condition was \((M=147.67)\), where read write learners were \((M=110.8)\), visual learners were \((M=148.33)\), and kinesthetic learners were \((M=128.4)\) in the auditory feedback condition. The mean NASA-rTLX scores for kinesthetic learners in the mixed modal feedback condition was \((M=122.5)\), read write learners were \((M=101.6)\), visual learners were \((M=170.5)\), and aural learners were \((M=139.83)\) in the mixed modal feedback condition. Finally, learners who tested as a two-letter combination VARK score, also known as a mixed learner, was examined relative to the mixed modal feedback condition, as all participants who tested as mixed
learners happened to be some variation combined with the kinesthetic learning style. The mean NASA-rTLX score for mixed learners in the mixed modal feedback condition was (M=198.15).

4.9 System Usability

The System usability scale (SUS) consisted of 10 questions. Scores were calculated by subtracting one from odd numbered scores, subtracting even numbered scores from 5, summing all new scores, and multiplying the total by 2.5. SUS scores are reported as percentiles. Scores above 68 are considered above average. Results for System Usability Scale scores showed a mean score of 79.12 (M= 79.12, SD=9.39). The distribution (N=34) of SUS scores can be seen below in Figure 6.

5. DISCUSSION

5.1 Response Times

When looking at the relationship between response time and the different feedback modalities delivered through the watch, an ANOVA revealed that the effect of feedback modality on response time was significant with a p value of .0001. However, post hoc analyses indicated that it was the auditory condition that had significantly slower response times on average at 4.58 seconds. While visual notifications delivered through the watch were synced with the timing of auditory notifications coming out of the speaker, it may have been noticeable enough to be distracting and potentially alter response times. The same effect is not seen in the mixed modal feedback condition where participants received auditory, haptic, and visual feedback. Controlled trials, where no
feedback was delivered, had the fastest response times on average at 2.66 seconds. One possible explanation for this may be that the act of simply attending to the watch is enough to add seconds to one’s response time. Average response times for the haptic feedback, mixed modal feedback, and visual feedback conditions were relatively similar with mean scores of 3.28 seconds, 3.24 seconds, 3.20 seconds respectively.

5.2 VARK

When analyzing VARK score results, 7 out of 34 participants tested as mixed learners. Mixed learners were identified when participants tied in any two learning style categories. Interestingly, 6 out of 7 of these learners tested as a combination with the kinesthetic learning style, with 1 testing as VA, 2 as VK, 2 as AK, and 2 as RK.

5.3 Number of Feedback Notifications Per Trial

Responses received in less than 2 seconds were removed from our sample, as these responses did not require feedback, which would have been delivered at the 2 second mark. The number of feedback notifications received per trial was examined across the various feedback conditions. The auditory feedback condition contained the lowest number of feedback notifications per trial on average (M = 2.72, SD= .42) with a p value of .0001. The haptic feedback condition received the highest number of feedback notifications per trial on average (M = 6.56, SD=.40).

When examining the distribution of feedback notifications over the 9 system failures relative to the different feedback modality conditions, there were no noticeable trends or
clustering as once suspected. This type of analysis was done in effort to see if performance increased by the end of the trial however this relationship was not seen, and all conditions had relatively even distributions of feedback over the two minute trial.

5.4 Situational Awareness

The auditory feedback condition had the highest mean score for SA at 24.38 while the mixed modal feedback condition had the lowest mean score at 22.15. On average situation awareness was the lowest when participants were exposed to the mixed modal feedback condition. The mean SA score was slightly higher for read write learners in the visual feedback condition with (M=22.6), compared to other learning styles in the visual condition such as visual learners (M=22.5), kinesthetic learners (M=22), and auditory learners (M=19.2). Situational awareness was higher on average when read write learners received their preferred method of instruction, which they received as written text in the visual feedback condition. The remaining learning styles did not show this relationship. Contrary to expectations, kinesthetic learners reported the lowest mean score for SA when they received their preferred method of instruction in the mixed modal feedback condition. One interesting finding was that the highest mean score for SA was reported (M=25.92) when participants who tested as mixed learners received mixed modal feedback. It is interesting to note that 6 out of 7 participants who tested as mixed learners were a combination with kinesthetic which VARK identifies as a learning style who prefers mixed modes of instructions.
5.5 Mental Workload

Of the 5 feedback conditions, the visual feedback condition reported the lowest mean NASA-rTLX score. On average, mental workload was the lowest when visual text was presented on the watch with a mean score of 113.87 and highest when participants received mixed modal feedback at 154.28. This means mental workload was the highest and situation awareness was the lowest in the mixed modal feedback trials. Controlled trials where no feedback was delivered scored similarly to the visual feedback condition with a mean mental workload score of 119.48. This could be because the mixed modal feedback overwhelmed sensory input channels where visual feedback and controlled trials were less obtrusive and demanded less attentional resources.

5.6 Learning Styles

Finally, in terms of learning style, means scores for mental workload were relatively lower when read write learners received their preferred method of instruction, through written text in the visual feedback condition. Mean mental workload scores were higher for aural learners when they received their preferred mode of instruction in the auditory feedback condition. The highest reported mean score for mental workload was among mixed learners when they received mixed modal feedback.

Being a kinesthetic learner had a significant effect on response times with a $p$ value of 0.0053. Post hoc analyses indicated that for kinesthetic learner’s, response times in the auditory feedback condition were significantly slower than visual feedback, mixed modal feedback, and control conditions.
5.7 Usability

Results for SUS scores showed a mean score of 79.12 (M= 79.12, SD=9.39). Overall, system usability was rated high, as scores over the 68th percentile are considered above average.

6. CONCLUSION

6.1 Conclusion

While system usability was rated high, the nature of the interactions with the wearable device may have affected the results of this study. Overall, controlled trials that received no feedback had the fastest mean response times. This could be due to the fact that the act of physically attending to the watch took up too much time. Careful design considerations should be made when designing interactions with wearable devices and should account for this.

Auditory feedback had a significant impact on response times. However, response times were actually significantly slower than other conditions. Interestingly, participants received the least amount of feedback notifications in trials with auditory feedback.

Sample size and task difficulty may have also contributed to insignificant results. Future studies could explore larger populations and layer system failures, as this may provide the right level of complexity and be more representative of the real world.

In visual feedback trials participants reported the lowest mean scores for MWL of all feedback types. This may suggest that visual feedback has the ability to decrease MWL
or this may only mean that visual feedback was inherently less distracting relative to other feedback types in this study.

Participants reported the highest mean scores for SA in trials where they received auditory feedback. While response time performance did not align with subjective ratings of situation awareness, this may suggest that participants perceived auditory feedback as more favorable, and future studies should include questionnaires to capture subjective ratings of feedback modalities and their perceived usefulness.

MWL was the highest and SA was the lowest in the mixed modal feedback trials. The effect of the mixed modal feedback may have overwhelmed sensory input channels which could have contributed to these results.

In terms of learning styles, read write learners on average showed both increased SA and decreased MWL when they were presented with their preferred instructional style, through written text in the visual condition. The same relationship was not seen for kinesthetic learners in the mixed modal feedback trials or aural learners in the auditory feedback trials as expected.

In conclusion, these results provide little evidence for learning styles or meshing theory. This study was designed as a first step to begin to examine the design of training programs for complex systems and how feedback can be optimized to improve operator performance in such settings. Future studies should continue to iterate on and refine design concepts and feedback interactions with smart watches like the one used in this study so that design recommendations can be made for these specific types of human computer interactions.
7. REFERENCES


8. APPENDIX

8.1 Consent Form page 1/3

Subject Informed Consent Document

Adaptive Feedback Interface

Investigator(s) name & address: Thomas Merrell, Jr., B.S.
Layla Aklan, B.S.
Subhashini Ganapathy, Ph.D.
Department of Biomedical, Industrial, and Human Factors Engineering.

Site(s) where study is to be conducted: Wright State University

Phone number for subjects to call for questions: (937)775-5044

Introduction and Background Information

You are invited to participate in a research study. The study is being conducted by Thomas W. Merrell, Jr., B.S., Layla Aklan, B.S., and Subhashini Ganapathy, Ph.D. in the Department of Biomedical, Industrial, and Human Factors Engineering. Approximately 50 subjects will be invited to participate.

Purpose

The purpose of this study is to develop a better understanding of how people respond to feedback interfaces for training purposes. These specific feedback interfaces will provide information to aid training the participants to correctly accomplish the tasks. In this study, participants will be using the Multi-Attribute Task Battery developed by NASA to be a flight simulator. A key goal of this study is to determine specifically if there is a relationship between performance in a multi-tasking situation, preferred learning style and the feedback method.

Procedures

In this study, a total of 50 participants age 18 to 50 inclusive are anticipated in this study. Participants will undergo a brief training session to learn how to operate the controls of the MATB simulator and the smart watch application. Once the training session has concluded participants will follow a randomized run order to reduce the effects of training bias. Following each trial, a NASA-TLX and SART survey will be administered to evaluate the mental workload and situational awareness. After completing all the trials, participants will receive a system usability scale survey to evaluate the usability of each of the feedback methods.

During testing participants will be alone in a testing room with the investigators observing through a one-way mirror in an adjacent room, with the screen mirrored for observation as well. All participants will be seated at a desk during testing. In all the experiment is expected to run on average 2 hours.

Potential Risks

There are risks associated with the MATB simulator. Those risk(s) is/are possible mental stress. This can be easily avoided or resolved by ending the scenario. There is also the potential loss of

V 205.19.2018

Page 1 of 3
8.2 Consent Form page 2/3

Confidentiality, in which identifying information may be exposed. This risk will be minimized by storing any material that may identify participants in a swipe access only laboratory in a locked filing cabinet that only the investigators and advisor will have access to, as well as storing any electronic files on a password protected computer in the same laboratory.

Benefits

The possible benefits of this study include furthering the participant’s understanding of how they best learn, the same can be said for humanity, if a relationship can be found that will allow people to be better trained for complex tasks.

The information collected may not benefit you directly. The information learned in this study may be helpful to others.

Compensation

You will be paid (or receive) $50 ($10 for applying, $40 for being accepted into the study) for your time, inconvenience, or expenses while you are in this study.

Confidentiality

Total privacy cannot be guaranteed. We will protect your privacy to the extent permitted by law. If the results from this study are published, your name will not be made public. Once your information leaves our institution, we cannot promise that others will keep it private.

Your information may be shared with the following:

- The Wright State IRB and Office of Research and Sponsored Programs
- Office for Human Research Protections (OHRP)

Security

The data that is collected will be stored both physically and electronically. The physical data will be kept in a card swipe access laboratory in a key locked file cabinet which only the PI and Faculty advisor can access. The electronic data will be stored in two ways, first on an external hard drive stored in the laboratory in the file cabinet, and second on a password protected server maintained by Wright State University which is only accessible by the PI.

Voluntary Participation

Taking part in this study is voluntary. You may choose not to take part at all. If you decide to be in this study, you may stop taking part at any time. If you decide not to be in this study or if you stop taking part at any time, you will not lose any benefits for which you may qualify.

Research Subject’s Rights, Questions, Concerns, and Complaints
8.3 Consent Form page 3/3

You may contact the principal investigator at merrill.8@wright.edu, aklan.4@wright.edu, or faculty advisor at subhashini.ganapathy@wright.edu.

If you have any questions about your rights as a study subject, questions, concerns or complaints, you may call the Wright State IRB Office (937) 775-4462. You may discuss any questions about your rights as a subject with a member of the IRB or staff. The IRB is an independent committee composed of members of the University community, staff of the institutions, as well as lay members of the community not connected with these institutions. The IRB has reviewed this study.

This paper tells you what will happen during the study if you choose to take part. Your signature means that this study has been discussed with you, that your questions have been answered, and that you will take part in the study. This informed consent document is not a contract. You are not giving up any legal rights by signing this informed consent document. You will be given a signed copy of this consent to keep for your records.

Signature of Investigator ______________________________ Date Signed ______________

Printed Subject Name ___________________________ Signature of Subject ___________________ Date ______________
Signed
The VARK Questionnaire (Version 7.8)

How Do I Learn Best?
Choose the answer which best explains your preference and circle the letter(s) next to it. Please circle more than one if a single answer does not match your perception. Leave blank any question that does not apply.

1. You are helping someone who wants to go to your airport, the center of town or railway station. You would:
   a. go with her.
   b. tell her the directions.
   c. write down the directions.
   d. draw, or show her a map, or give her a map.

2. A website has a video showing how to make a special graph. There is a person speaking, some lists and words describing what to do and some diagrams. You would learn most from:
   a. seeing the diagrams.
   b. listening.
   c. reading the words.
   d. watching the actions.

3. You are planning a vacation for a group. You want some feedback from them about the plan. You would:
   a. describe some of the highlights they will experience.
   b. use a map to show them the places.
   c. give them a copy of the printed itinerary.
   d. phone, text or email them.

4. You are going to cook something as a special treat. You would:
   a. cook something you know without the need for instructions.
   b. ask friends for suggestions.
   c. look on the Internet or in some cookbooks for ideas from the pictures.
   d. use a good recipe.

5. A group of tourists want to learn about the parks or wildlife reserves in your area. You would:
   a. talk about, or arrange a talk for them about parks or wildlife reserves.
   b. show them maps and internet pictures.
   c. take them to a park or wildlife reserve and walk with them.
   d. give them a book or pamphlets about the parks or wildlife reserves.

6. You are about to purchase a digital camera or mobile phone. Other than price, what would most influence your decision?
   a. Trying or testing it.
   b. Reading the details or checking its features online.
   c. It is a modern design and looks good.
   d. The salesperson telling me about its features.

7. Remember a time when you learned how to do something new. Avoid choosing a physical skill, eg, riding a bike. You learned best by:
   a. watching a demonstration.
   b. listening to somebody explaining it and asking questions.
   c. diagrams, maps, and charts - visual clues.
8.5 VARK Learning Style Inventory page 2/3

8. You have a problem with your heart. You would prefer that the doctor:
   a. gave you a something to read to explain what was wrong.
   b. used a plastic model to show what was wrong.
   c. described what was wrong.
   d. showed you a diagram of what was wrong.

9. You want to learn a new program, skill or game on a computer. You would:
   a. read the written instructions that came with the program.
   b. talk with people who know about the program.
   c. use the controls or keyboard.
   d. follow the diagrams in the book that came with it.

10. I like websites that have:
    a. things I can click on, shift or try.
    b. interesting design and visual features.
    c. interesting written descriptions, lists and explanations.
    d. audio channels where I can hear music, radio programs or interviews.

11. Other than price, what would most influence your decision to buy a new non-fiction book?
    a. The way it looks is appealing.
    b. Quickly reading parts of it.
    c. A friend talks about it and recommends it.
    d. It has real-life stories, experiences and examples.

12. You are using a book, CD or website to learn how to take photos with your new digital camera. You would like to have:
    a. a chance to ask questions and talk about the camera and its features.
    b. clear written instructions with lists and bullet points about what to do.
    c. diagrams showing the camera and what each part does.
    d. many examples of good and poor photos and how to improve them.

13. Do you prefer a teacher or a presenter who uses:
    a. demonstrations, models or practical sessions.
    b. question and answer, talk, group discussion, or guest speakers.
    c. handouts, books or readings.
    d. diagrams, charts or graphs.

14. You have finished a competition or test and would like some feedback. You would like to have feedback:
    a. using examples from what you have done.
    b. using a written description of your results.
    c. from somebody who talks it through with you.
    d. using graphs showing what you had achieved.

15. You are going to choose food at a restaurant or cafe. You would:
    a. choose something that you have had there before.
    b. listen to the waiter or ask friends to recommend choices.
    c. choose from the descriptions in the menu.
    d. look at what others are eating or look at pictures of each dish.

16. You have to make an important speech at a conference or special occasion. You would:
    a. make diagrams or get graphs to help explain things.
    b. write a few key words and practice saying your speech over and over.
    c. write out your speech and learn from reading it over several times.
    d. gather many examples and stories to make the talk real and practical.

8.6 VARK Learning Style Inventory Scoring page 3/3
The VARK Questionnaire Scoring Chart

Use the following scoring chart to find the VARK category that each of your answers corresponds to. Circle the letters that correspond to your answers.

e.g. If you answered b and c for question 3, circle V and R in the question 3 row.

<table>
<thead>
<tr>
<th>Question</th>
<th>a category</th>
<th>b category</th>
<th>c category</th>
<th>d category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>K</td>
<td>A</td>
<td>R</td>
<td>V</td>
</tr>
<tr>
<td>2</td>
<td>V</td>
<td>A</td>
<td>R</td>
<td>K</td>
</tr>
<tr>
<td>3</td>
<td>K</td>
<td>V</td>
<td>R</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>K</td>
<td>A</td>
<td>V</td>
<td>R</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>V</td>
<td>K</td>
<td>R</td>
</tr>
<tr>
<td>6</td>
<td>K</td>
<td>R</td>
<td>V</td>
<td>A</td>
</tr>
<tr>
<td>7</td>
<td>K</td>
<td>A</td>
<td>V</td>
<td>R</td>
</tr>
<tr>
<td>8</td>
<td>R</td>
<td>K</td>
<td>A</td>
<td>V</td>
</tr>
<tr>
<td>9</td>
<td>R</td>
<td>A</td>
<td>K</td>
<td>V</td>
</tr>
<tr>
<td>10</td>
<td>K</td>
<td>V</td>
<td>R</td>
<td>A</td>
</tr>
<tr>
<td>11</td>
<td>V</td>
<td>R</td>
<td>A</td>
<td>K</td>
</tr>
<tr>
<td>12</td>
<td>A</td>
<td>R</td>
<td>V</td>
<td>K</td>
</tr>
<tr>
<td>13</td>
<td>K</td>
<td>A</td>
<td>R</td>
<td>V</td>
</tr>
<tr>
<td>14</td>
<td>K</td>
<td>R</td>
<td>A</td>
<td>V</td>
</tr>
<tr>
<td>15</td>
<td>K</td>
<td>A</td>
<td>R</td>
<td>V</td>
</tr>
<tr>
<td>16</td>
<td>V</td>
<td>A</td>
<td>R</td>
<td>K</td>
</tr>
</tbody>
</table>

Calculating your scores

Count the number of each of the VARK letters you have circled to get your score for each VARK category.

Total number of V circles = _____________
Total number of A circles = _____________
Total number of R circles = _____________
Total number of K circles = _____________

8.7 Situation Awareness Rating Technique (SART) page 1/2

Situation Awareness Rating Technique

Participant ID: ___________________
Trial #: ________________

Instability of Situation
How changeable is the situation? Is the situation highly unstable and likely to change suddenly (high) or is it very stable and straightforward (low)?

Complexity of Situation
How complicated is the situation? Is it complex with many interrelated components (high) or is it simple and straightforward (low)?

Variability of Situation
How many variables are changing within the situation? Are there a large number of factors varying (high) or are there few variables changing (low)?

Arousal
How aroused are you in the situation? Are you alert and ready for activity (high) or do you have a low degree of alertness (low)?

Concentration of Attention
How much are you concentrating on the situation? Are you concentrating on many aspects of the situation (high) or focused on only one (low)?

Division of attention

How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (high) or focused on only one (low)?

Spare Mental Capacity

How much mental capacity of you have to spare in the situation? Do you have sufficient capacity to attend to many variables (high) or nothing to spare at all (low)?

Information Quantity
How much information have you gained about the situation? Have you received and understood a great deal of knowledge (high) or very little (low)?

Familiarity with Situation

How familiar are you with the situation? Do you have a great deal of relevant experience (high) or is it a new situation (low)?

Information Quality

How valuable is the information received from the situation? Is the quality of the information very good (high) or is the quality very poor (low)?
NASA-TLX

Participant ID: _______________
Trial #: _______________
8.9 System Usability Scale page 1/1

<table>
<thead>
<tr>
<th>1. I think that I would like to use this system frequently</th>
<th>1 2 3 4 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. I found the system unnecessarily complex</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>3. I thought the system was easy to use</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>4. I think that I would need the support of a technical person to be able to use this system</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>5. I found the various functions in this system were well integrated</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>6. I thought there was too much inconsistency in this system</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>7. I would imagine that most people would learn to use this system very quickly</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>8. I found the system very cumbersome to use</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>9. I felt very confident using the system</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>10. I needed to learn a lot of things before I could get going with this system</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>