An Exploration of Tactile Warning Design Based on Perceived Urgency

by

Yeti Li

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Applied Science

in

Systems Design Engineering

Waterloo, Ontario, Canada, 2013

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

When there is information overload on the visual modality, another system of warnings must be adopted to prevent potential risks—tactile warning systems present a viable alternative. Building on work on design approaches for auditory warning systems that match appropriate warnings to the severity of risk, this thesis presents an approach to design tactile warnings based on perceived urgency. To do this, I use a subjective rating technique. I performed three experiments to demonstrate this approach.

Our research approach uses subjective rating technique to evaluate perceived urgency. Three experiments were conducted to design tactile warnings with a tactile interface developed by attaching a grid of tactors on a vest. In Experiment 1 and 2, I evaluated perceived urgency of several warning designs with three important parameters of tactile warnings with subjective rating. In Experiment 3 I examined one warning design in the context of flight simulation.

The results of Experiment 1 and 2 showed that participants can discriminate between all levels of perceived urgency from most warning parameters. In Experiment 3, the results showed that selected warning design was correctly mapped with the severity of most events. The findings suggest that tactile warnings based on perceived urgency can be a possible approach, but further studies will be required to evaluate different parameters of tactile warnings.

Acknowledgements

I wish to thank my supervisor Professor Catherine M. Burns for bringing me into the field of cognitive human factors engineering. Her assistance, patience, and advice kept my research on the correct track.

I would also like to thank my readers, Professor Carolyn MacGregor and Professor John Zelek, for their comments, support, and guidance on my research.

Robert Arrabito and Geoffrey Ho of Defense Research and Development Canada, Toronto are especially thanked for supporting me during my MASc studies.

I want to express my gratefulness to my colleagues in the Advanced Interface Design Lab and external partners, Plinio P. Morita, for helping me all the time as the lab manager; Wayne Giang and Vivek Kant, for helping with the early hypothesis of this study; Garry Dunn of Trellis Consulting, for providing software consulting services; Anson Ho for help with experimental design; David Yoon for running the experiments; and Leila S. Rezai for help with data analysis. A special thanks to. Special thanks to Katie Webb for taking the time to read through this thesis and providing valuable comments.

Thanks to all the student volunteers at the University of Waterloo, who spent their valuable time participating in the experiments.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

Human errors have gained unprecedented attention, especially in complex systems. Warnings are designed and implemented in human-machine interfaces to provide cues and details of potential risks and hazards. They allow people to detect risks and take proactive actions to avoid these dangers. Human errors can thus be reduced with the help of such warnings. However, on many occasions the presence of too much information can result in information overload, because of the limited decision making resources possessed by humans. The sources of information are usually varied, making it difficult for people to distinguish useful information, especially under severe work and time pressure. In many situations, warnings are unnoticed or the severity of them is misjudged. In these cases, people do not respond to warnings correctly so no appropriate corrective actions are taken. Visual interfaces are commonly used to present information and it is in this modality that information overload occurs frequently. This potential for overload indicates that new solutions should be considered in the design of warnings.

One solution is to present warnings to other senses. Many studies focus on the auditory interface as a substitute modality for presentation of warnings (Suied, Susini, & McAdams, 2008; Jang, 2007; Edworthy, Loxley, & Dennis, 1991). Tactile interfaces, however, have only recently become of interest in the study of warnings. The development of computer technologies that allow information presentation on through vibrotactors (Mori, Tanaka, & Kaneko, 2011; Hoggan & Brewster, 2007; Gemperle, Ota, & Siewiorek, 2001), has created an interest in the development of tactile interfaces and the ability of this technology to use the human body as a carrier for warning systems.

Another problem in the design of warnings is that warnings do not always match the severity of the hazard. This issue becomes more critical if warnings are presented when information overload is occurring. The misjudgement of the degree of risk can lead to incorrect decisions and serious incidents. One solution for this problem is to design warnings that match the appropriate levels of urgency to the severity of the hazards. This warning design approach based on perceived urgency (Edworthy, 1994; Hellier, Edworthy, Weedon, Walters, & Adams, 2002) was initially proposed for auditory warnings. The level of urgency for each warning was assessed through a subjective rating. The idea of such an urgency rating came from the finding that the acoustic parameters of auditory stimuli have intrinsic connections to levels of urgency (Edworthy et al., 1991; Patterson, 1982). Therefore, specific warnings are able to map to the urgency of risks. This approach has also been extended to the design of visual warnings (Rashid & Wogalter, 1997; Chapanis, 1994). To our knowledge, this approach has not been extended to the design of tactile warnings, before this work.

This thesis is an exploration of these principles as applied to the design of tactile warnings. Though there are many research questions that require further study, we focused on two major topics: a design approach to extend perceived urgency mappings to tactile interfaces and a set of experiments to demonstrate that tactile warnings can be designed with specific mappings that create different perceived levels of urgency.

There are many frameworks and principles for warning design and many works are available as empirical guidance for their development and evaluation. There are some studies on perceived urgency for auditory and visual warnings that quantify the effects of warnings and map warnings to the appropriate levels of risk. For tactile warning design, many techniques exist, but no study directs test the research on perceived urgency within this modality. I developed a framework based on auditory warning design to develop and test mappings for urgency in tactile warnings with the four-phase design approach for auditory warning design of Patterson (1982). I used an adaptation of Edworthy (1991) and Arrabito's (2009) urgency assessment methods to assess our designs. My approach also involves mapping tactile warnings to different degrees of risks and hazards.

This thesis focuses on the implementation of the tactile design approach (Chapter 3). In preliminary work, we developed a wearable tactile display to present tactile information (Chapter 4). The first experiment manipulated three parameters of tactile stimuli and developed a set of tactile warnings. In this experiment, an urgency rating task was completed by 10 participants. Participants simultaneously completed a secondary evaluation, in which a simple validation of tactile warnings was performed. The second experiment was identical in design and procedure to the first experiment; however, this experiment performed further study on one of the three parameters of tactile stimuli that was found to be the most promising in the first experiment. In our last experiment, three tactile warnings which were selected from our first two experiments were tested in the context of flight simulation and participants rated perceived urgency of the tactile warnings during five different events in the simulation.

This thesis is organized as follows: Chapter 2 provides a review of two techniques that contribute to our study: warning design based on perceived urgency and tactile interface design. Chapter 3 presents our design approach for tactile warning design based on perceived urgency. Chapter 4 discusses the settings, procedures, and results of experiments 1 and 2. Chapter 5 explains the design and results of Experiment 3. In Chapter 6 I draw conclusions and discuss limitations and ideas for future work.

Chapter 2
Background

In this chapter, I review relevant research to lay a foundation for tactile warning design based on perceived urgency. First we focus on the origin, methodologies, and applications of perceived urgency. This discussion provides a guideline to refine the value of auditory warning studies on perceived urgency, and constitutes a basis for a similar approach to tactile warning design. After this, we discuss studies on tactile interface design. This is the starting point to understanding some of the major characteristics of tactile interfaces and parameters of tactile stimuli.

2.1 Warning Design Based on Perceived Urgency

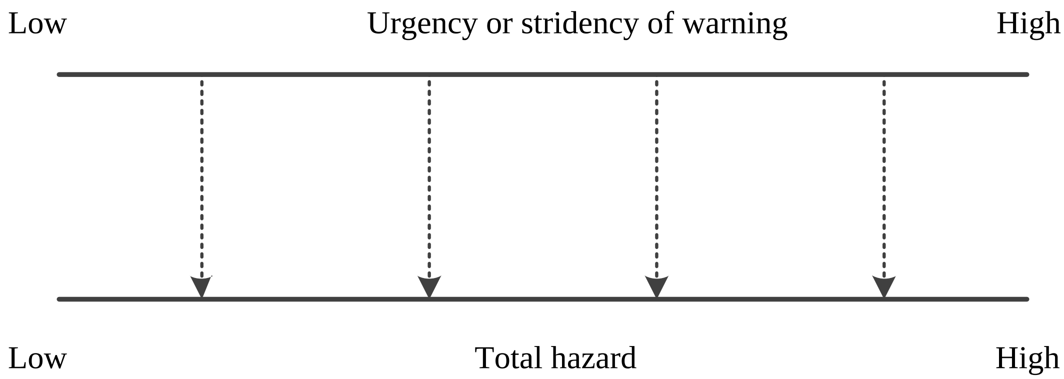
In this section, the discussion follows several stages of warning design based on perceived urgency. First, we present the concept of perceived urgency and the motivation for urgency-based warning studies. A detailed review of the auditory warning design approach on perceived urgency follows. This review provides the approach for warning stimuli design and urgency assessment techniques. Finally, we briefly discuss the limitations of this approach.

2.1.1 Perceived Urgency

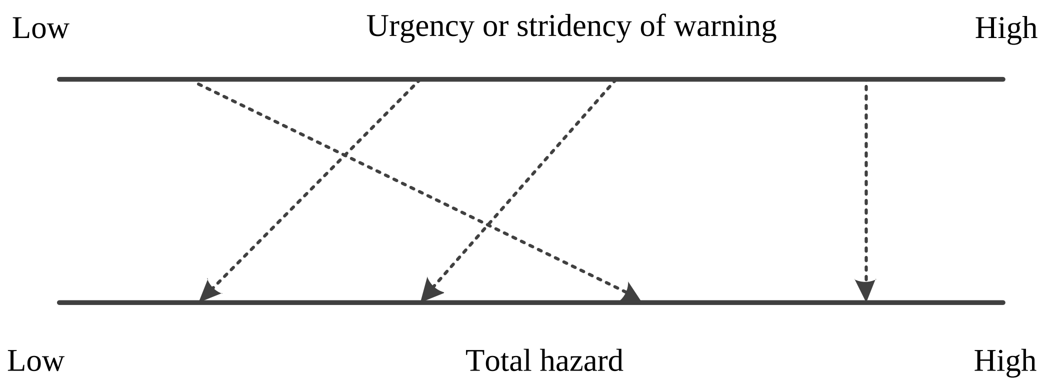
Warnings are designed to alert people that "something bad is happening.” Urgency is one of the important features of warning designs, which allows people to differentiate risks and hazards based on their severity. In emergencies, people should be notified of risks by their priority, especially in cases where there is a threat of information overload. To do this, the use of urgency is preferred by many warning researchers as a method to encode the priority of risks when using warning signals. For example, in one experiment (Chih-Yuan, Nikolic, & Sarter,

2001) participants were asked to complete an interruption cueing task in three modalities: visual, auditory, and tactile. Participants were divided into two groups. In the basic group, participants were only notified about the presence of interruption tasks. In the other group, the urgency of the tasks was also presented. The results showed that there was a significant improvement in performance on the task when urgency information was presented. This example indicates the effectiveness of encoding urgency in warnings.

Ideally, warnings should be correctly associated with the priority of actual situations (Figure 1-a). In practice however, the designed warnings sometimes inappropriately reflect the hazardousness of a situation, meaning a high urgency warning is assigned to a non-urgent situation (Edworthy, 1994; Figure 1-b). Thus, the term "urgency mapping" was raised in the study of auditory warning design (Edworthy & Adams, 1996). There are two types of urgency, as classified in auditory warning studies (Burt, Bartolome, Burdette, & Comstock, 1995; Edworthy et al., 1991b). In auditory warning design, perceived (psychoacoustic) urgency is determined by people who perceive warning signals. Situational urgency is assigned with the severity of the current situation. Thus, situational urgency is correlated with "desired" actions for the situation. There are some examples of people working in a high-workload environment confused when identifying the meaning of alarms, especially when these alarms are unreliable (Bliss & Dunn, 2000; Momtahan & Tansley, 1989). One of the reasons for this is that people perceive a warning, but match this warning with an incorrect risk or hazard. Thus, the mismatch in urgency mapping between perceived urgency and situational urgency occurs.



a) Correct mapping of perceived urgency



b) Inappropriate mapping of perceived urgency

Figure 1: Examples of urgency mapping (adapted from Edworthy & Adams, 1996, p. 7).

It is important to study perceived urgency in designing warnings, as it is essential for effective urgency mapping. There are two major components to each warning (Edworthy & Adams, 1996): an iconic component and an informational component (Figure 2). The iconic component catches immediate attention and causes quick responses. For example, a lighted exit sign alerts people to an emergency. The informational component can carry some additional information (usually associated with recommendations of further actions; for example, an emergency exit sign also directs a safe way in hazard such as fire). However the informational component is dependent on the essential correctness of the information. If the information is wrong, people may take disastrous actions. For example, if the exit sign directs to a dead end in emergency, lives will be lost. Perceived urgency is an essential feature in the iconic component of the warnings (Edworthy & Adams, 1996). Thus, the level of perceived urgency should be considered in warning design. To summarize, it is very helpful if warnings can be designed using mappings that match the level of perceived urgency and that perceived urgency is appropriately mapped with the severity of hazards.

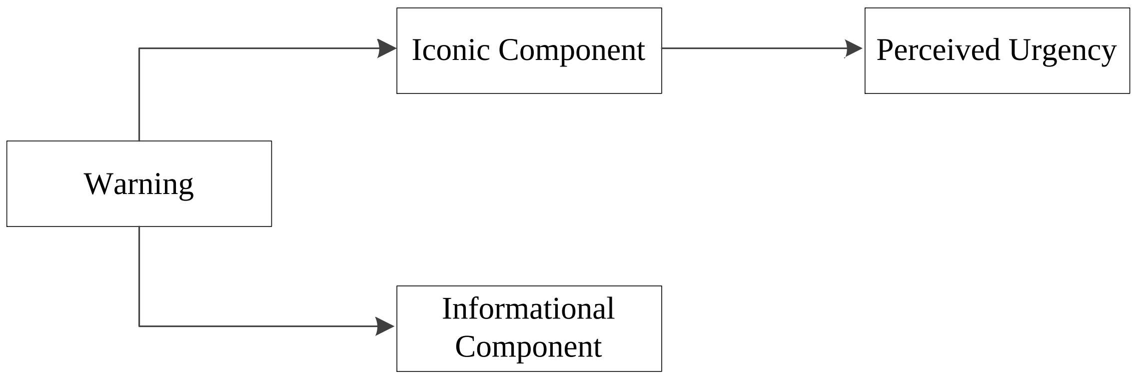


Figure 2: Perceived urgency as the one of the iconic components.

2.1.2 Auditory Warning Design Approach Based on Perceived Urgency

Auditory warnings have been designed to have specific mappings to trigger perceived urgency. Studies show that perceived urgency is significantly influenced by the acoustic parameters of warnings (Edworthy & Hellier, 2006; Wiese & Lee, 2004; Edworthy et al., 1991; Patterson, 1982), such as fundamental frequency, overall length of the auditory signal, and intensity of sound. Some experimental studies have examined the effects of individual sound parameters (i.e. fundamental frequency). Subsequent studies have shown that some parameters contribute greater weight to perceived urgency than others (Hellier, Edworthy, & Dennis, 1993), such as pulse format, pulse level, and inter-pulse interval (Haas & Casali, 1995). Later there was a trend to concentrate on verbal warnings, including effects of voice warnings (Park & Jang, 1999), verbal semantics and acoustics (Hellier et al., 2002), and acoustic versus non-acoustic parameters (Jang, 2007).

An early study (Patterson, 1982) proposed a four-stage strategy for the development of auditory warning stimuli (Figure 3). In the first step, an appropriate sound level is selected. Then the design focuses on pulse design. The pulse of sound indicates a short stimulus, lasting approximately 100 to 300ms and is determined by several temporal characteristics of the sound. A burst is designed by repeating a pulse while differing pitches and amplitudes. Normally, it lasts 2s and works as a complete auditory warning. For the last step, some inter-bursts of silence is added. By combining different parameters, warning designers can develop different warnings. This principle provides guidance for the development of auditory stimuli.

Step 1. Appropriate sound level
Step 2. Pulse design (fundamental frequency, harmonic series, etc.)
Step 3. Combine pulses into a burst
Step 4. Insert an inter-bursts silence

Figure 3: Strategy to develop auditory warning stimuli (adapted from Patterson, 1982).

The next step in warning design is to consider the effectiveness of that warning. A subjective rating technique was developed by Edworthy et al. (1991) to assess the effects of these parameters on the perceived urgency of auditory warnings. Many experiments have examined the contribution of one particular parameter on perceived urgency of auditory warnings; for example: non-vocal warnings (Guillaume, Pellieux, Chastres, & Drake, 2003), warnings in hospital operating rooms (Mondor & Finley, 2003), and speech warnings (Hellier et al., 2002). A rating scale has been used for urgency ranking in many auditory warning experiments (Edworthy et al., 1991; Arrabito, Mondor, & Kent, 2004). The initial study (Edworthy et al., 1991) was conducted with a scale of 0 to 100 for perceived urgency rating. For this study, 16 subjects, with ages ranging from 18 to 40 years, participated in three experiments of warnings. Three types of acoustic parameters were tested in the experiment. A total of 13 auditory warnings were designed by combining different sound parameters.

Table 1: Example of urgency rating (partial results of Experiment 1; Edworthy et al., 1991).

Fundamental Frequency (Hz)	Harmonic Series	Envelope Shape	Ratings (Rank)
150	Regular	Standard	38.4 (7)
150	Regular	Slow offset	23.9 (13)
150	Regular	Slow onset	36.3 (10)
150	10% Irregular	Standard	52.0 (4)
150	10% Irregular	Slow offset	28.0 (11)
150	10% Irregular	Slow onset	53.5 (3)
150	Random	Standard	70.2 (1)
530	Regular	Standard	50.2 (5)
530	Regular	Slow offset	27.6 (12)
530	Regular	Slow onset	36.6 (9)
530	10% Irregular	Standard	61.8 (2)
530	10% Irregular	Slow offset	36.8 (8)
530	10% Irregular	Slow onset	47.1 (6)

Table 1 shows an example of the stimuli tested and results of rating. Warnings can be ranked by their ratings. Thus, the rating technique indicates the perceived urgency of each warning. This approach has shown that it is possible to achieve a clear level of perceived urgency by manipulating acoustic parameters.

2.1.3 Limitations of Perceived Urgency

Though a correct urgency mapping can be achieved from the subjective rating technique, there are a few concerns regarding the effectiveness of warnings. The presence of warnings should lead to a change in human behaviour. Whether this change occurs determines the effectiveness of warnings. Information on warning compliance comes from early studies on compliant behaviours, such as during evacuations (for example: Perry, Greene, & Lindell, 1980; Ikeda, 1982).

Unfortunately, there is still not much evidence on how perceived urgency can improve warning compliance. The arousal effect suggests that a compliant behaviour can be provoked with specific temporal sound parameters in both nonverbal warnings (Suied et al., 2008) and verbal warnings (Jang, 2007). Response time, could be considered indicative of arousal and is also an interesting measure that can be measured in real time. A high level of perceived urgency has been demonstrated to be correlated with short response times. Response time is typically measured in experiments by testing warnings of different urgency. Some such studies have shown that a higher perceived urgency may result in a shorter response time (Arrabito et al., 2004; Haas & Casali, 1995). Beyond this finding,the relationship between the level of urgency of a warning and compliant behaviour is still unclear. The true test of compliance would require the examination of warnings in the field. This was outside of the scope of thesis but could be considered for further work.

2.2 Tactile Interface Design

The skin of the human body can be used as a substitute interface for presenting information (Veen & Van Erp, 2001), especially when the information overload in other modalities occurs. Many studies have built tactile interfaces to provide spatial and localization information via the skin (Van Erp & Self, 2008; Gallace, Tan, & Spence, 2007; Van Erp, Veen, Jansen, & Dobbins, 2005). The human part which is used as the tactile interface is varied. Examples include: wrist belts (Van Erp, Veen, Jansen, & Dobbins, 2005), waistbands (Cholewiak, Brill, & Schwab, 2004) and vests (Van Erp & Veen, 2003; Wu, Zhang, Yan, Liu, & Song, 2012). To transfer information (including warnings) with the tactile modality, careful signal design is essential. Studies on tactile signals provide many recommendations on manipulating tactile signals. Jones et al. suggest the best range of frequencies for the tactile sense which is critical for a detection or localization task (2008). Brown et al. indicate that introducing complex rhythms and waveforms are more effective in designing effective tactile signals than adjusting parameters such as frequency and duration (2005). It has been proven that manipulating parameters of tactile stimuli convey differences in the performance for these tasks. However, there is little research on how these parameters can individually or systematically effect perceived urgency on tactile interfaces.

As with auditory warnings, tactile warnings require correctly mapping urgency with the severity of risks. Burns et al. (2011) presented a systematic approach for using cues to map information with visual, auditory, and tactile modalities. It has been demonstrated that there are limitations on the types of information that can be presented to the tactile modality. For example, the tactile modality can be used to encode one or two dimensional analogical information; however, it is not a suitable modality to utilize continuous information. This is because of the presence of tactile adaptation (Nafe & Wagoner, 1941). This adaptation occurs when constant vibration to the same location leads to a significant decrease in sensitivity.

2.3 Conclusion

In summary, the motivation to develop tactile warnings based on perceived urgency comes from the noticeable advantage of this approach in developing warnings that map well to the severity of situations. Over the last two decades, this approach has become more prevalent in the study of auditory warning designs. Another motivating factor is the growing interest in, and achievements of tactile interface design. Existing studies have provided examples of methods of tactile information presentation and mapping, but have not correlated tactile design parameters with perceived urgency. This review of the literature on auditory warning design and tactile interface design, based on perceived urgency, has set the theoretical basis for our approach to tactile warning design.

Chapter 3
Approach for Tactile Warning Design Based on Perceived Urgency

This chapter proposes an approach to tactile warning design. Herein, we summarize a four-step approach to warning design based on perceived urgency and provide a detailed plan for each step within the scope of tactile warning design. The guidelines begin with techniques to develop different warnings. After that, the subjective rating technique is implemented to examine the urgency level of each warning. An appropriate urgency mapping is then developed for certain situations. In the last step, we include a secondary evaluation for participant's response to warnings. Together, this presents a preliminary view of the effectiveness of designed warnings.

3.1 Stimuli Design

This discussion of design procedure (Figure 4) starts from the stage of stimuli design, in which the warning interface is established and warning signals are developed. This stage involves many modality-specific characteristics and capabilities.

We build on the four-step cycle for stimuli design developed by Patterson (1982). His original design cycle begins with a sound level, which described the fundamental attribute of an auditory signal. As auditory warnings were usually transferred to people via a headphone or a speaker, the fundamental sound level was decided by the playback equipment and designed volume. In tactile signals, tactors are used in many studies to present warning signals. Tactors are small transducers that provide haptic feedback to human skin. In our study, they were activated to transfer tactile stimuli to human skin. Some similar attributes in tactile interface are gain and frequency. As discussed in Chapter 2, tactors were used on wrist, belt, or vest to compose an interface. At this stage, designers should determine a tactile interface to present the warnings. Therefore Patterson's first step in stimuli design was kept for our tactile design framework.

The second step in auditory stimuli design is manipulating other acoustic parameters to develop a pulse. As discussed in Chapter 2, a pulse is a short (200 – 300ms) sound. In this case, it reflects specific temporal characteristics of sound (Edworthy et al., 1991). People should be able to discriminate different auditory warnings, so people can memorize these warnings as clearly as possible. We made a similar attempt on tactile warning design. One of the major shortcomings in tactile warning research is that the correlation between specific tactile parameters and perceived urgency is unclear. At this stage, researchers and designers should test multiple parameters in developing a tactile warning (for example, the amplitude and the duration of the signal). By combining different parameters, a set of tactile warnings can be developed and tested for mapping to perceived urgency.

The third step in auditory warning design is modality specific. A burst is a longer and melodic signal that consists of several pulses and takes approximately 2s (Edworthy et al., 1991). One of the reasons to build an auditory burst is to improve the quality of sound, as an auditory pulse is too short to memorize. To improve the quality of tactile signal however, warning signals can be adjusted based on forms of presentation. The spatial layout shows the type of tactile interface (for example; wrist, belt, or vest) and localization of signals. As we mentioned in Chapter 2, many studies showed that tactile interface can be used to present spatial and localization information (for example: Van Erp & Self, 2008; Gallace et al., 2007). However in our framework the spatial layout was used as a parameter and it was hypothesised to have some effects on perceived urgency.

The final step is to add a short period of silence between two tactile signals. Because of the adaptation phenomena of tactile warnings (Nafe & Wagoner, 1941), silence is essential in presenting tactile warnings.

3.2 Perceived Urgency Rating

The rating technique used in auditory warning design is extended here to tactile warning design. At this stage, experiments are planned that combine different parameters for testing. In particular, all the participants should be able to perceive tactile warnings well. Considering the physical variance, the tactile interface should be adaptable in size for all the participants to ensure good contact between the tactors and the skin. For example, use a resizable tactile belt for different people. However, because we have used spatial layout as a parameter, it is important to determine some reference locations on the interface. For example, the spine and the navel region of the abdomen were used as anatomical reference because people were more sensitive to detect vibrations on these points (Cholewiak et al., 2004). In this case, fitting of the tactile display is reliable with different participants.

3.3 Urgency Mapping

Mapping urgency for the design of tactile warnings is slightly different with such mapping in the auditory modality. In auditory warning design, urgency mapping is usually synthesized on auditory bursts rather than pulses to improve the quality of warning. Therefore, the urgency of auditory warnings is used as a temporal characteristic. (Arrabito et al., 2004; Edworthy et al., 1991) Tactile warnings, however, should not be presented consistently because of the presence of tactile adaptation (Nafe & Wagoner, 1941). Chapter 5 gives an example to map urgency of tactile warnings to critical events, which change gradually. In the example, a short period of silence was added between two tactile stimuli.

3.4 Other Measures

As discussed in Chapter 2, when designing auditory warnings, response time was measured in some studies (Arrabito, 2004; Haas & Casali, 1995). In Experiment 1, a simple, secondary task is planned in which participants are asked to identify a tactile warning as quickly and as accurately as possible. The accuracy and response time will be recorded to see if there is any significance between different tactile warnings. This provides some extra data for the preliminary evaluation of tactile warning design based on perceived urgency.

3.5 Conclusion

Building on approaches to designing auditory warnings based on perceived urgency, in this chapter we propose an approach to tactile warning design based on perceived urgency. Step 1 (stimuli design) builds on Patterson's four-step cycle (1982). Steps 2 (perceived urgency rating) and 4 (other measures) also build on pre-existing work on auditory warning design. Step 3 (urgency mapping) is based on modality specific experimentation. In order to stay within a reasonable scope for discussion however, for this thesis we limit our approach to the tactile warning design based on perceived urgency.

Figure 4: Approach to tactile warning design based on perceived urgency.

Chapter 4

Experiment 1 and 2: Tactile Warning Design and Perceived Urgency Rating

This chapter follows the design approach in Chapter 3 and discusses how to develop tactile warnings. Two experiments were completed to examine three important parameters of tactile warning design that could improve subjective perceived urgency rating. In each experiment, the same rating scale was used. Both experiments were based on a tactile display that was made of a grid of tractors. The tactile display was mounted on a vest worn by participants in the experiments. The two experiments were identical in general experimental settings and procedures. The first experiment focused on three parameters of tactile warning: activation level, activation type and layout type. The activation level refers to the vertical position of the tactile activation that occurred on the vest. Both activation type and layout type indicate the temporal variations of tactile display presentation.

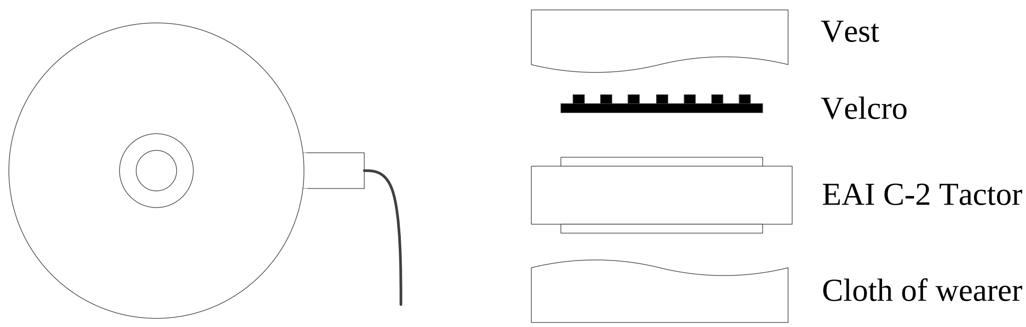
Based on the results of Experiment 1, the most promising type of activation was examined further in the second experiment. Since this activation type could be explored in two different layouts, both layout alternatives were tested.

4.1 The Tactile Display

EAI[1] electromechanical tactors were used to build the tactile display. The C-2 tactor is a light weight tactor which has been widely used already in many tactile interface studies (for examples: Brown et al. 2006, Van Erp et al. 2001). A vest was used as a supportive frame to hold the vibrotactors in contact with the user. Three sizes of nylon vests (small, medium and large) were provided for the tactile display so a wide variety of user sizes could be fitted appropriately. The inner side of the vest was rough, covered by Velcro loops, which allowed tactors, with corresponding Velcro hooks on them, to be attached. The Velcro allowed tactors to stay firmly on the vest and vibrate without resistance. By using Velcro the tactors could be repositioned to tailor the fit more precisely to the users. Experiment participants were instructed to wear thin clothes, a t-shirt was the suggestion, which allowed the tactor vibrations to be transmitted to the skin. Figure 5 shows how a tactor is attached between the wearer and the vest.

---

[1] EAI (Engineering Acoustics, Inc.) website: http://eaiinfo.com/home.htm



a) A tactor b) A tactor attached on the vest

Figure 5: Attach a tactor to the vest.

A total of 9 tactors were placed on the vest in a square grid of 3 by 3 tactors. Figure 6-a illustrates the grid. The centre column of tactors provides a reference axis which was positioned on the spinal column of the participant. Past research suggests that the wearer may not be able to feel an acuity difference between two tactors on the torso if tactors are placed quite near (2-3 cm) (Van Erp & Veen, 2003). Therefore, both vertical and horizontal distances between adjacent tactors were set firmly at 4 cm. The grid was initially placed 4 cm away from the upper edge of the vest (Figure 6-b) for each wearer as recommended. However, the grid was repositioned along the spinal column if vibration on any of the tactors could not be felt well by the participant.

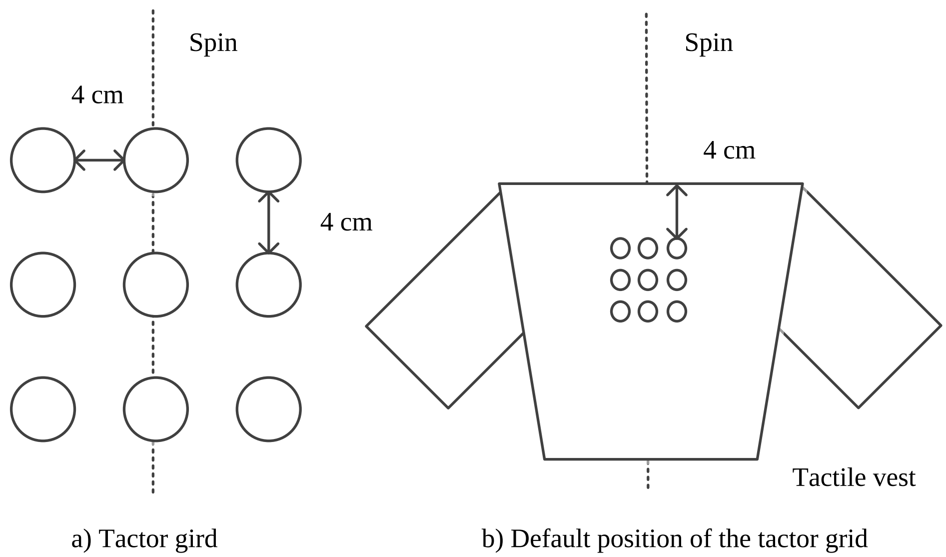


Figure 6: The grid of tactors.

After the grid of tactors was built, all the tactors were connected to a Tactor Control Unit (also provided by EAI). The control unit provides a maximum of 24 output channels, which means that up to 24 tactors can be activated simultaneously. A software package was also provided along with the hardware, including a visualized development tool named "TActionWriter". Specified by this tool, a TAction carries some major parameters of the stimuli, such as frequency and which tactor(s) is being called by the control unit. I made one TAction for each tactile stimulus as designed in the experimental paradigm. More details of the stimuli are discussed in the next section.

Another tool in the software package provided by EAI was an ActiveX library named "TActionReader". The library was called in the experimental program, which was made with Microsoft Visual C++. When a tactile stimulus was being played, the TActionReader read a

TAction from file and sent commands and tactile parameters to the tactor control unit. The unit activated specific tactor(s) and the participant can receive the tactile stimuli from the vibration of tactors in their back. The tactile display was used in both experiments.

4.2 Experiment 1

4.2.1 Design

Following the tactile warning design approach in Chapter 3, I started my work of building tactile stimuli from the fundamental vibration level. First, a fundamental frequency was set for all stimuli. A past study by Jones & Safter (2008) demonstrated that the optimal frequency for tactile sense reception on human skin should be set within 150 to 300 Hz. Human sensitivity for detecting vibration signals is close to a parabolic curve (Verrillo, 1966), and the frequency at the lowest threshold sensitivity is around 250 Hz. For this reason, in our experiment the fundamental frequency was set at 250 Hz for all the tactile stimuli.

The total duration of each stimulus was set at 200ms. According to our tactile design approach, a 20ms period of silence was added before the tactile activation in each design to limit the vibration strength. The activation duration for each trial was designed as 180ms so the ratio of time for activation was 20% (duty cycle). The duty cycle was designed this way to allow our tactile display to avoid pressure adaptation, which would reduce the skin sensitivity after a long and constant tactile stimulus (Nafe & Wagoner, 1941). Participants were asked to rate perceived urgency right after each trial (stimulus). The presentation of tactile display was stopped until participants finished rating, although they were told to indicate their rating as soon as possible. Tactile stimuli were not played continually. Thus, the period of silence before the tactile activation cannot be tested as a tactile design parameter. However, the period of silence was essential when these warnings were later implemented in real situations for urgency mapping, especially when a constant warning was required (Chapter 5).

Therefore, three tactile parameters were examined at this stage. By combining different levels of these parameters, we developed the set of tactile warnings for urgency rating in this experiment. In the following sections, we discuss these parameters and the hypothesis of the influence of these parameters on perceived urgency.

Activation levels

We hypothesised that a vibration occurred higher (closer to the shoulders) on a participant's back may convey higher level of urgency. If so, the column of tactors may be used to present warnings of different severities. Thus, from the topmost row to the bottommost row of tactors, the activation was named as Level 3, Level 2 and Level 1 (Y-axis, Figure 7).

Activation types

We hypothesized that an increase in the number of activated tactors may produce higher perceived urgency in subjective rating. To examine this, we tested two different activation types, as shown in (Figure 7): Target and FullBar. In Target design, only the top most tactor(s) was activated as the activation level increased. Thus, the urgency of this design was encoded only with the spatial location of vibration. For FullBar, however, all the tactors underneath were also activated. In this design, perceived urgency could be influenced by both the spatial activation of the top tactor, but also by the number of tactors activated at each urgency level. We would hypothesize, therefore, that the FullBar design should provide a stronger mapping of perceived urgency than the Target design.

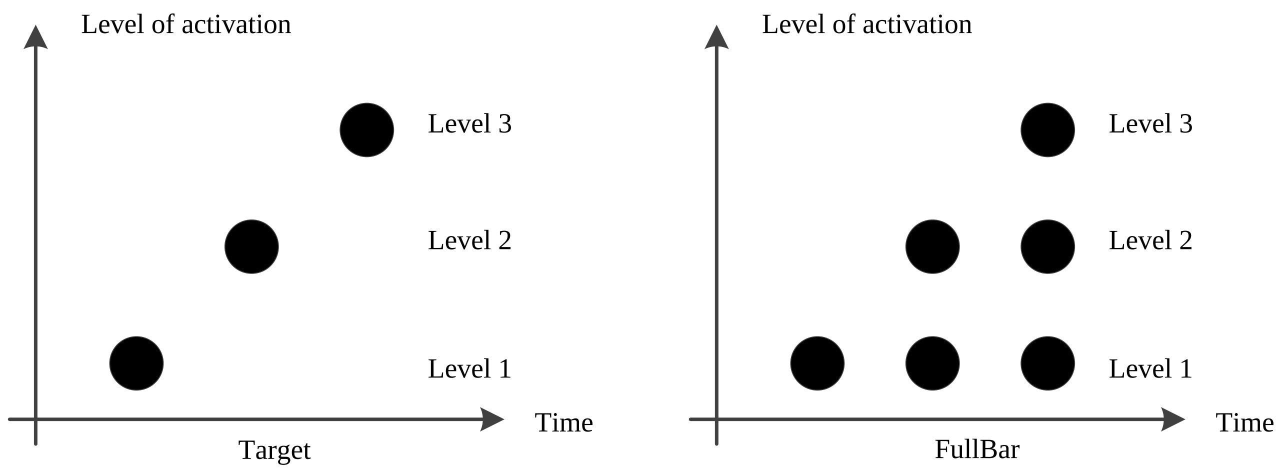


Figure 7: Activation types in Experiment 1.

Layout types

Another difference in the number of activated tactors occurred when varying layout type. Two layout types were designed, and they were different in the number of columns used on the tactor grid (Figure 8). In the OneColumn design, only the middle (spinal) column of tactors was used for activations, while in the TwoColumn design only the left and right column were activated to present tactile stimuli. We hypothesized that the TwoColumn design should result in stronger urgency mappings, because more tactors are activated in each warning. However, it should be noted that the spine is a very strong referent in tactile display (Cholewiak et al., 2004) and a sensitive region on the back, which could interfere with the OneColumn design.

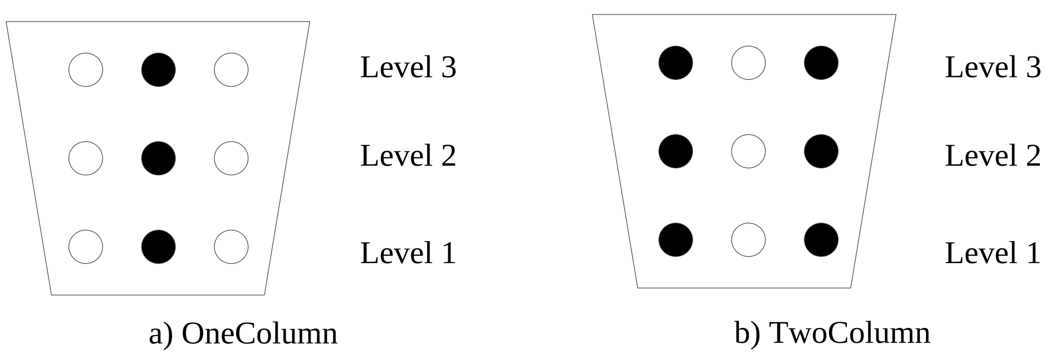


Figure 8: Layout types in Experiment 1.

4.2.2 Stimuli

The term "design" was used as the combination of activation type × layout type. The activation level was designed to have three grades, and it was hypothesised to be persistent with all designs. Thus, we hoped the activation level would map with the degree of severity in each design, and we could compare the effectiveness of all designs. Therefore, the development of trials included all possible combinations of the parameters. A "configuration" in this experiment was defined as the combination of activation type × layout type × activation level. Each configuration was presented as a different tactile stimulus. Perceived urgency was rated for each trial by participants.

Table 2: Configurations in Experiment 1

Activation type (2)	Layout type (2)	Activation level (3)
Target	OneColumn	Level 1
FullBar	TwoColumn	Level 2
		Level 3

Table 2 shows that a total of 12 configurations were built for Experiment 1. One TAction file was made by TActionWriter for each configuration. We first defined the fundamental frequency of the tactile stimulus in the TActionWriter. After that, we determined which tactor(s) to be activated for each TAction. Finally, the activation duration and the period of silence were set. The TActions were later compiled and ported to a library, where the stimuli can be loaded from the file and played in the experimental program.

4.2.3 Hypothesis

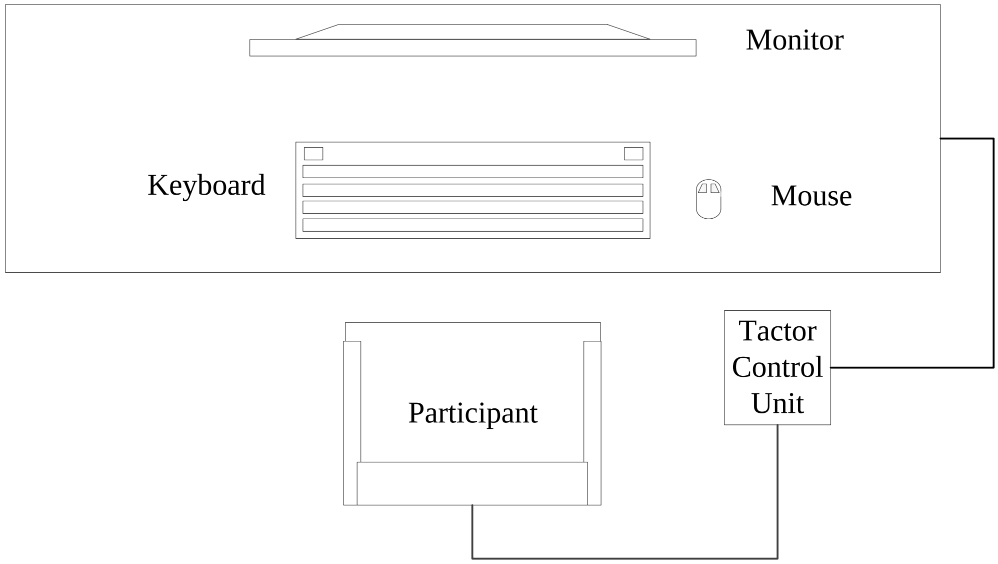
In conclusion, we summarize our hypotheses as follows:

1. In general, more activated tactors should result in higher perceived urgency. This means that FullBar warnings should be perceived more urgent than Target warnings because more tactors are being activated. This also suggests that TwoColumn warnings should be perceived more urgent than OneColumn warnings because more tactors are being activated.

2. Higher (closer to the shoulders) activation levels should result in higher perceived urgency. This means that Level 3 warnings should be perceived more urgent than Level 2 warnings; Level 2 warnings should be perceived as more urgent than Level 1 warnings. Thus, we hypothesize that activation level can be used to present three degrees of urgency for each design (activation type $\times$ layout type).

4.2.4 Apparatus

The experiment took place in an experiment room at University of Waterloo. All sessions were arranged in quiet hours to minimize the disturbance of extraneous sounds. The color of interior schemes and furnishings were brown, and the ambient lights were dimmed during the experiment. Figure 9 shows the equipment setting in the experiment room.



.

Figure 9: Experimental set up for Experiment 1.

The same setup was used in was used in both Experiment 1 and 2. All the experimental guidance and information were presented through a 22-inch liquid crystal display monitor. Participants were asked to respond to the tasks by using the number pad of a keyboard in front of the monitor. The mouse was only used in the training section. Inspired by tactile localization performance may differ from head orientations finding (Ho & Spence, 2007), participants were asked to remain looking at the monitor throughout the experiment.

4.2.5 Procedure

Participants were told to wear a T-shirt before coming to do the experiment in order to keep the clothing layer between the tactors and the skin thin. At the start of the experiment, participants were given an information letter and a consent form. The information letter described the goal and procedures of this experiment. If the participants agreed to participate in the experiment, they were told to sign the consent form. After that, participants were asked to remove wristwatches and turn off all electronic devices such as cellphones. Participants were able to try all three sizes of vests (small, medium and large) and pick the most suitable size. Then, they were asked to sit on the chair in front of the monitor.

The experiment started with a training section. First, participants were introduced to the three parameters in the tactile warning design. After that, each participant was given a familiarization task to acquaint the participant with all the trials in the experiment. Trial, in this case, refers to the presentation of a particular configuration of tactors. Each trial was presented on the tactor grid in a Target or FullBar sequence. During each trial, a window was shown on the monitor similar to Figure 10. The window displayed the name of the currently played trial and three buttons underneath. Participants were not required to memorize onscreen names of the trials, but they were asked to remember spatial locations of the three activation levels (for example, Level 1 represented a tactile warning on the bottommost row of tactors). Since this was a familiarization task, the experimenter used the mouse to replay the current trial, proceed to the next trial or return to the previously played trials. Participants were asked if they could feel all the vibrating tactors. If not, the experimenter repositioned the tactor grid along the spinal column and restarted the familiarization task to make sure all tactors were perceived.

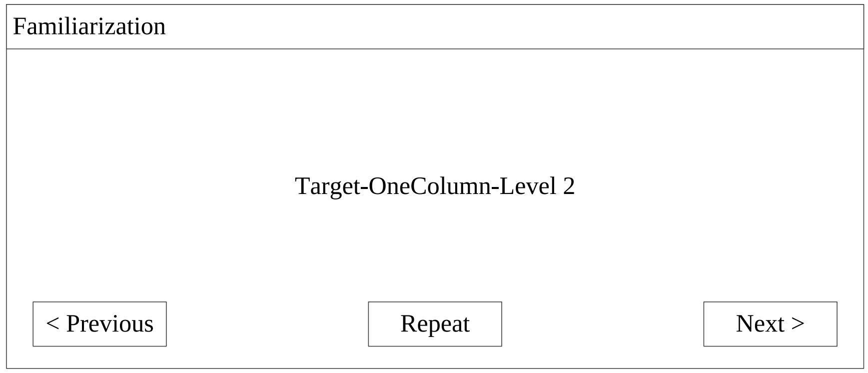


Figure 10: Prompt window for familiarization in Experiment 1.

When the training session was completed, the participant started the discriminability task. In the discriminability task, all the trials were presented in a randomized sequence. After each trial was played, a window was shown on the monitor (Figure 11). Participants were asked to choose the level of activation as quickly and as accurately as possible, by pressing the key 1, 2 or 3 on the number keyboard. To confirm the selection, the participants pressed the enter key. Each tactor configuration was replicated five times at some point in the randomized sequence. Thus, 60 trials were played for each participant.

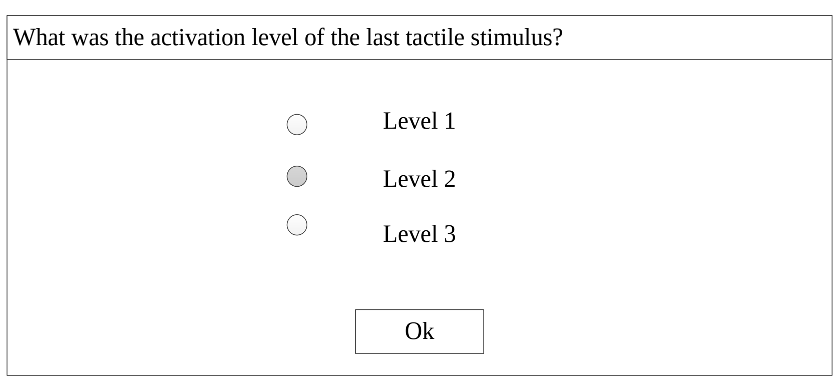


Figure 11: Prompt window for the discriminability task in Experiment 1.

After completing the discriminability task, participants were asked to rate the level of perceived urgency for each tactile configuration. Each trial in the urgency rating task, presented a randomized tactile configuration. After the presentation of each trial, a rating slider was shown on the monitor similar to Figure 12. The rating scale of 1 to 100 was based on Edworthy (1991) and Arrabito's (2004) urgency assessment method. Participants moved the mouse to drag the slider, and click the "Ok" button to complete a rating. Each tactile configuration was presented five randomized times throughout the urgency task.

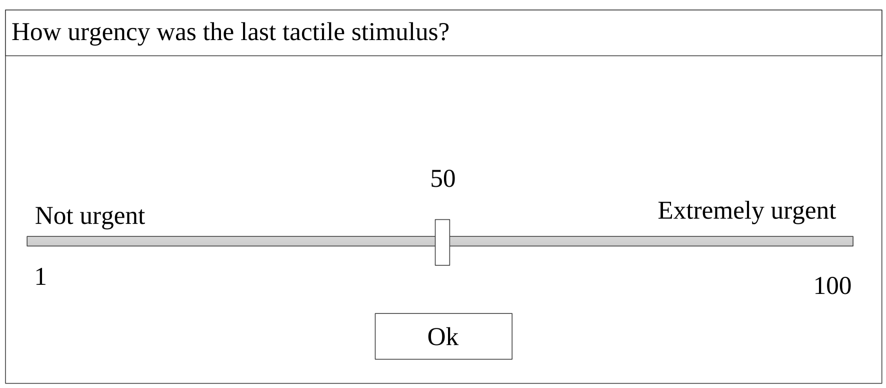


Figure 12: Prompt window for urgency rating task in Experiment 1.

Following all experimental tasks, participants were asked to fill out a questionnaire on their overall preferences between the designs. As well the questionnaire asked the participants to comment on whether they experienced any annoyance from the tactile designs. This was included because long duration of tactile stimuli can be irritating. Kaaresoja & Linjama (2005) suggested that the vibration duration of tactile stimuli should be between 50 and 200ms. The sample questionnaire is listed in Appendix B. 1. The whole experiment took approximately 1 hour to complete.

4.2.6 Participants

Ten participants took part in the experiment. All participants were undergraduates and graduate students recruited from the University of Waterloo. Six participants were female and four participants were male. Table 3 is a summary of the completion of Experiment 1. All ten participants chose the medium-size vest, and all of them completed the experiment without adjusting the position of the tactors. All the participants completed the discriminability task, the urgency rating task and the questionnaire.

Table 3: Completion of Experiment 1.

Task	Number of participants
Completed discriminability task	10
Completed urgency rating task	10
Completed the questionnaire	10

4.2.7 Results

Data analysis of the urgency rating task, discriminability task and the questionnaire are discussed respectively as follows:

Urgency Rating

The urgency rating scores were recorded by the experimental program and exported to a log file for each participant. The log files were manually aggregated to a summary table in Microsoft Excel. Following the subjective rating technique in auditory warning design (Arrabito et al., 2004; Edworthy et al., 1991), we calculated the mean rating of five replicates per configuration. Afterwards, the mean ratings were normalized to accommodate for individual scaling differences with a calculation in Appendix A.1. First, the consistency of rating between trials was examined by the Friedman test, Kendall's $W' = 0.692,\ x^2(11) = 76.140, p < 0.0005$, which was highly significant. The assumptions for normality ($p > 0.05$ for all groups) and homogeneity ($F(11, 109) = 0.613, p = 0.814 > 0.05$) were met as prerequisites for the evaluation on variances.

A factorial, three-way ANOVA of activation type (Target versus FullBar) × layout type (OneColumn versus TwoColumn) × level of activation (Level 1 versus Level 2 versus Level 3) was conducted on normalized mean rating. Results showed there were highly significant main effects of all independent variables on normalized mean rating: activation level, $F(1, 60) = 14.32, p < .0005$; activation type, $F(1, 60) = 44.81, p < .0005$; layout type, $F(1, 60) = 20.97, p < .0005$. No statistically significant three-way activation type × layout type ×

activation level interaction was found, $F(2, 109) = 0.69, p = .502$. There was a statistically significant activation level $\times$ layout type interaction, $F(2, 109) = 7.64, p = .001$. No other significant interactions were observed.

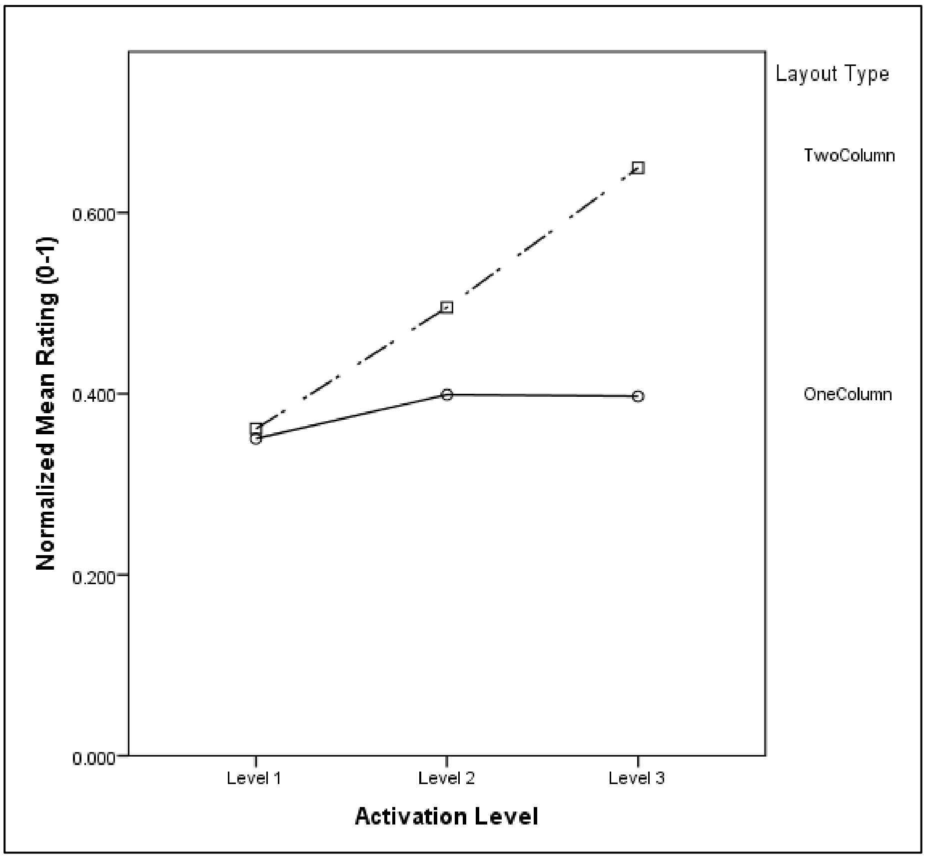


Figure 13: Line chart of activation level × layout type in Experiment 1 (from SPSS)

Figure 13 shows the two-way activation level $\times$ layout type interaction. In the TwoColumn designs, normalized mean rating was significantly higher at activation Level 3 ($M = 0.650, SD = 0.030$) when compared to Level 2 ($M = 0.495, SD = 0.030$), $p = 0.002 <$

0.05, and Level 1 ($M = 0.358, SD = 0.029$), $p = 0.006 < 0.05$. However, in the OneColumn designs, no significant difference of normalized mean rating was found between activation levels.

Figure 14 shows the normalized mean rating of each configuration (activation type × layout type × level of activation) in the experiment

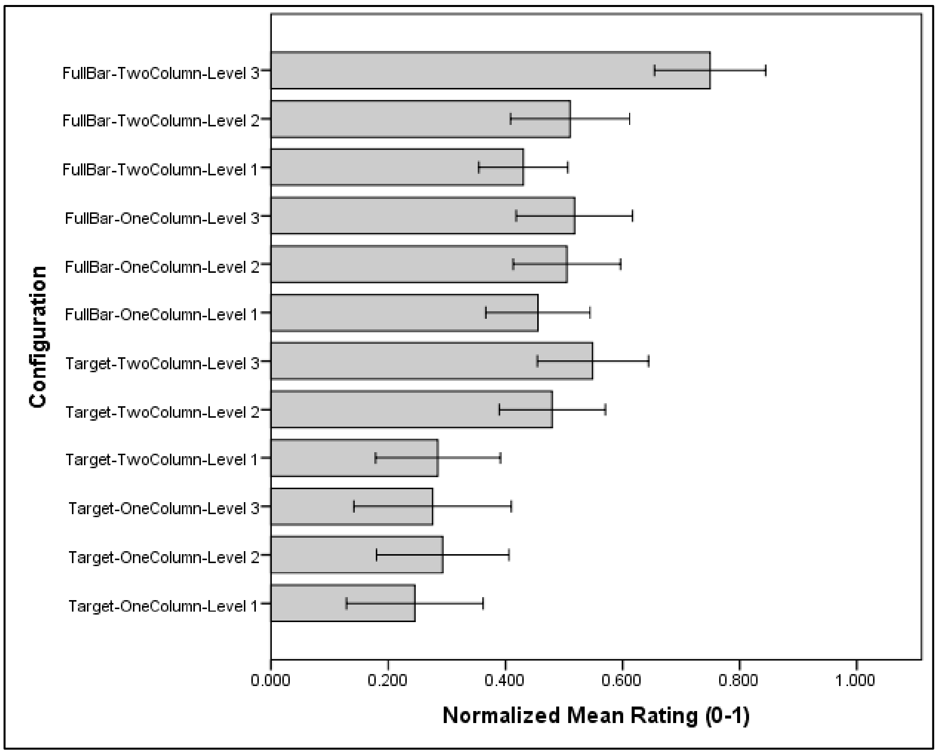


Figure 14: Clustered bar chart with confidence intervals in Experiment 1 (from SPSS).

Discriminability - Accuracy Rate

In the discriminability task, the activation levels selected by participants for each design (activation type × layout type) were recorded. The accuracy rate for each design was the number of accurate selections divided by the number of replications (five). The accuracy rate data passed the tests for normality in each design ($p > .05$). Levene's Test indicated, there was homogeneity of all variances ($p = .212$). A two-way ANOVA was conducted on accuracy rate with the activation type (Target and FullBar) and layout type (OneColumn and TwoColumn) as independent variables. No statistically significant main effect or interaction effect was found of either activation type or layout type on the accuracy rate ($p > .05$). Descriptive statistics showed that the overall accuracy rate was 0.630. Figure 15 shows the mean accuracy rate in each of the four designs: 0.687 for the Target-OneColumn warnings, 0.573 for the Target-TwoColumn warnings, 0.627 for the FullBar-OneColumn design, and 0.633 for the FullBar-TwoColumn warnings.

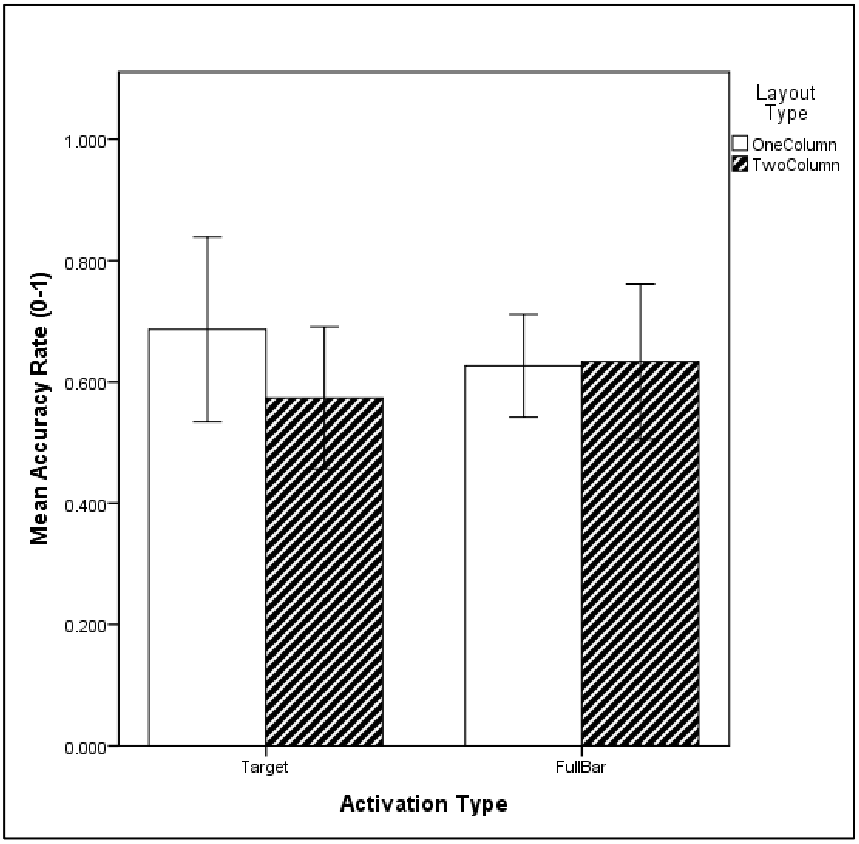


Figure 15: Accuracy rate of four designs in Experiment 1 (from SPSS).

Discriminability - Response Time to Correct Selections

The response time to select activation levels in the discriminability task was recorded in milliseconds in the experimental program. The results were calculated by the mean of all five replications per design (Figure 16). Only correct selections were included in the calculation of response time. All the distributions of response time were normal($p > .05$), and the assumption of homogeneity was met ($p = .657$). A two-way ANOVA was conducted on the mean response time of correct selections with the activation type × layout type. No statistically significant main effects were found for activation type or layout type, $p > .05$.

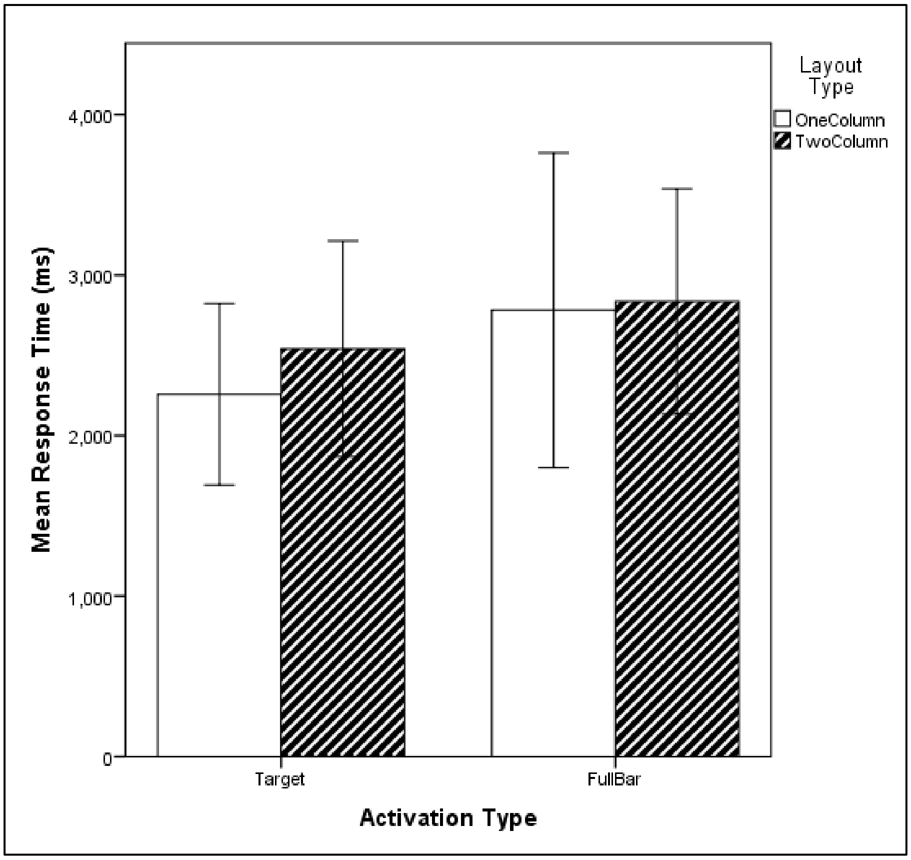


Figure 16: Response time of correct selections of four designs in Experiment 1 (from SPSS).

Questionnaire – Preference

Participants were asked for their overall preference in the four designs, namely Target-OneColumn, Target-TwoColumn, FullBar-OneColumn and FullBar-TwoColumn. A Pearson Chi-square test for association was conducted on preference for design. There was a statistically significant association between design and preference, $\chi^2 = 21.867, \mathrm{p} < .0005$. Descriptive statistics in Table 4 show that most participants preferred the FullBar-TwoColumn design.

Table 4: Preference in design in Experiment 1.

Design	Participants' Preference
Target-OneColumn	0%
Target-TwoColumn	10%
FullBar-OneColumn	10%
FullBar-TwoColumn	80%

Participants were asked to circle the most preferred activation type. Pearson Chi-square test on preference in design showed that no statistically significant association was found between activation type and preference, $\chi^2 = .800, p = .371$. From Table 5 we know that the preference for Target and FullBar was almost equally divided.

Table 5: Preference for activation type in Experiment 1.

Design	Participants' Preference
Target	40%
FullBar	60%

The said question was asked for participant's preference in two layout types. Results of a Pearson Chi-square test performed on preference in layout type showed, there was a statistically significant interaction between design and preference, $\chi^2 = 12.800, p < .0005$. Descriptive statistics showed almost all participants preferred the TwoColumn design (Table 6).

Table 6: Preference for layout type in Experiment 1.

Design	Participants' Preference
OneColumn	10%
TwoColumn	90%

Questionnaire – Understandability

The next metric was perceived understandability of the four designs. Participants were asked to circle the design which seemed easiest to understand. A Pearson Chi-square test showed that there was a significant association between design and perceived understandability, $\chi 2(1) = 22.933, p < .0005$. Table 7 reports that most participants believed that the FullBar-TwoColumn was easiest to understand. No Target designs were rated easiest to understand by participants.

Table 7: Understandability for design in Experiment 1.

Design	Participants' Preference
Target-OneColumn	0%
Target-TwoColumn	20%
FullBar-OneColumn	0%
FullBar-TwoColumn	80%

Participants were asked to choose which activation type was perceived to be easier to understand between the Target and FullBar designs. A Pearson Chi-square test was conducted on understandability for activation type, $\chi2(1) = 7.200, p = .023$, which was statistically significant. Most participants believed the FullBar design was easier to understand than Target, as shown in Table 8.

Table 8: Understandability for activation type in Experiment 1.

Design	Participants' Preference
Target	20%
FullBar	80%

Participants were asked which layout type seemed easier to understand, between the one column and two column layouts. All participants believed TwoColumn was easier to understand (Table 9). The result of Pearson Chi-square test was highly significant, $\chi2(1) = 20.000, p < .0005$.

Table 9: Understandability for layout type in Experiment 1.

Design	Participants' Preference
OneColumn	0%
TwoColumn	100%

Questionnaire – Annoyance

Participants were asked to rank the perceived annoyance of the four designs on scale from “the least annoying” (coded as 1) to “extremely annoying” (coded as 4). Results of a Friedman test suggested that there was no significant difference in annoyance between the four designs, $\chi2(3) = 3.154,\ p = .369$. Table 10 shows the mean annoyance score and standard deviation per design. Finally participants were asked if any of the designs were “too annoying to use”. None of the designs were rated as “too annoying to use”.

Table 10: Mean annoyance score of design in Experiment 1.

Design	M	SD
Target-OneColumn	2.778	1.302
Target-TwoColumn	3.000	0.866
FullBar-OneColumn	2.333	0.866
FullBar-TwoColumn	2.000	1.225

4.2.8 Discussion

Experiment 1 provided the first test of our preliminary designs as well as experience with our measurement approach. The tactile display was mounted on a vest with a grid of 3×3 tactors. Three tactile warning design parameters were proposed and combined to build tactile warnings for the experimental tasks. In the data analysis, a number of metrics were measured in the urgency rating task and discriminability task. In the urgency rating task, the result of Friedman's test was highly significant. The result shows that participants rated all of the tactile trials consistently, and they understood which designs were perceived to be more urgent than others. The finding supported perceived urgency as an important component of warning.

The results of the factorial ANOVA showed some effects of the three parameters (activation type, layout type and level of activation) to the perceived urgency of tactile stimuli. The finding was consistent with prior auditory warning studies that showed that the parameters of stimuli may have some effects on the perceived urgency of warnings. Our second finding came from the comparisons of levels in activation type and layout type. For activation type, specifically, the FullBar design was rated more urgent than the Target design. For layout type, the TwoColumn design was rated higher than the one column design. The results supported our hypothesis that more activated tactors result in higher perceived urgency (Figure 17).

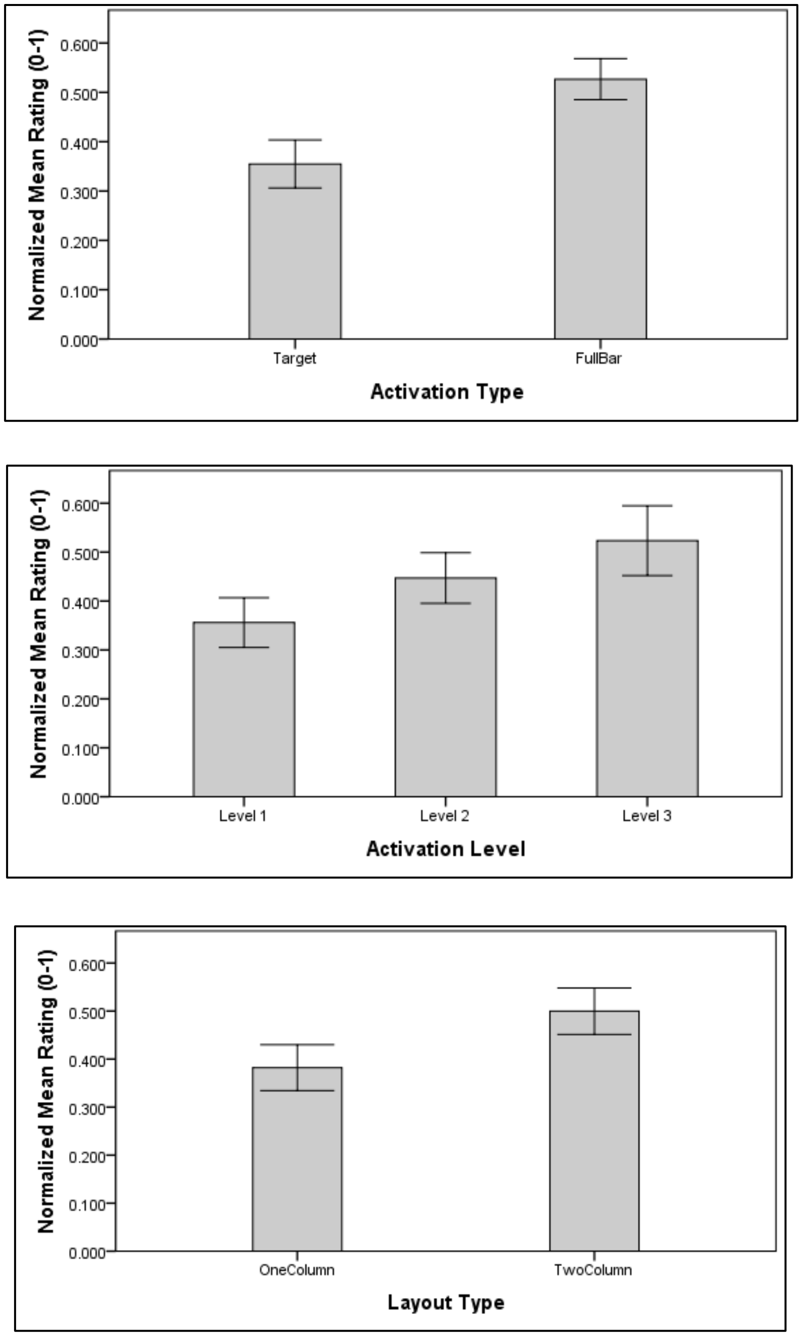


Figure 17: Parameters contribute to perceived urgency of Experiment 1 (from SPSS).

Figure 17 also suggests that activation level has some effect on perceived urgency. The highest rated levels were, in order, Level 3, Level 2 and Level 1. The two-way activation level ×

layout type interaction gave us a view that in the TwoColumn design, activated tactors on the bottommost row of the vest (Level 1) produced less urgency than those in the middle row (Level 2) and the topmost row of the vest (Level 3). This result partially supported our hypothesis that we can encode tactile warnings of three levels of urgency on activation levels.

Some comments from the questionnaire showed that it was hard to discriminate warnings at Level 1 and Level 2, but it was relatively easy to identify the highest level (Level 3). This result suggested that even in some conditions the level of urgency can be encoded to activation level, user's discriminability of these levels may still need further studies.

The effect of activation type is clear: a FullBar design produces higher urgency than a Target design and that support our hypothesis that activating more tactors increases the perception of urgency. No interaction between activation type and activation level was found, showing that either of the two types may be promising as a good design. However, results of the discriminability task revealed that the accuracy rate of Target designs was not significantly lower than FullBar designs in the discriminability task. One drawback of the Target design was found on the questionnaire that it was hard to differentiate between activation levels, which may have caused some confusion in urgency rating. To that extent, the FullBar design was a relatively more coherent design, which was also testified by results of understandability of activation type from the questionnaire.

The effect of layout type on perceived urgency seemed consistent: TwoColumn warnings resulted in higher urgency ratings than OneColumn warnings. Another finding comes from the activation type × layout type interaction that activation level has an effect on perceived urgency (the higher activation level conveys higher urgency) in the TwoColumn design but not in the OneColumn design (Table 11). A possible reason was that in OneColumn warnings the activation levels were not detected correctly, as only the tactors on the spinal column were activated in OneColumn designs. This was surprisinglydifferent from result of Cholewiak et al. work (2004) as we discussed in the experimental design. It seemed that the sensitivity on spine (OneColumn design) was not perceived more urgent than the increased column of tactors (TwoColumn design).

Results of preference and understandability of layout types also support that the TwoColumn design is a more consistent design.

Table 11: Normalized mean rating of activation level × layout type in Experiment 1

	Layout Type	
Activation Level	OneColumn	TwoColumn
Level 1	0.350	0.361
Level 2	0.399	0.495
Level 3	0.397	0.650

No significant association between the parameters and performance was discovered in the discriminability task. The overall accuracy rate was 0.630. In response time to correctly indicate activation level, no significant effect was found in either of the two parameters. These results were contrary to the Jones et al. (2009) experiment, where a vibrotactile pattern recognition task presented on the torso showed 98% correct accuracy levels. One possible reason for our lower accuracy results could be the lack of context in our experiment. It is also possible that the activation duration of each tactile stimulus was short (180ms). In the Summers et al. (1997) experiment, it was found that human discrimination improves as the duration of tactile stimuli increases from 80 to 320ms. In the Jones et al. (2009) experiment the activation duration of each tactile stimulus was 500ms. Therefore, in our experiment some of the trials may have been ignored (activation duration of each was 180ms) as participant's ability to detect the tactile stimuli was not strong, thereby impacting results on the discriminability task.

The results of Experiment 1 suggested that some of the proposed tactile design parameters may influence perceived urgency. The results also suggest that activation level can be used to present three levels of perceived urgency. The participant's preference for activation type and layout type was strong - the FullBar design produced less perceived confusion than the Target design, and the TwoColumn design was perceived to be more understandable at all activation levels than the OneColumn design. No designs received perceived annoyance ratings that would limit them from future consideration. In summary, the FullBar-TwoColumn design was recommended at this stage (Figure 18).

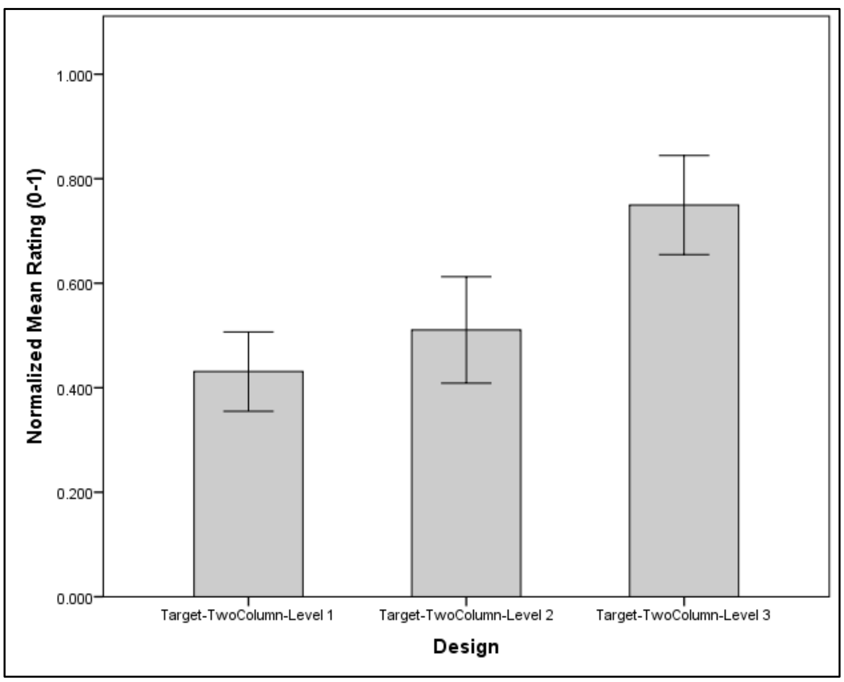


Figure 18: Clustered bar chart of FullBar-TwoColumn design of Experiment 1 (from SPSS).

One limitation was found in the participant group of this experiment. The recruitment of participants reflected the young demographic at University of Waterloo. Past studies (Goble, Collins, & Cholewiak, 1996; have suggested that human sensitivity to vibration may decrease when the age increases. Thus, the results of this experiment would require further verifications before generalizing to the general population.

A limitation of this experiment was that the designs were tested in the absence of a realistic context. For this reason, later in this thesis the designs are tested in the context of a flight simulation in Experiment 3 (Chapter 5). However a strong finding from this experiment suggested that the number of tactors activated was important in improving urgency, discrimation and user preference. In considering these results we realized that a further design possibility was available using the same tactor grid, a ThreeColumn design and that this design merited exploration. The ThreeColumn design is explored in Experiment 2.

4.3 Experiment 2

4.3.1 Design

Results of the first experiment revealed that more activated tactors result in higher perceived urgency. As TwoColumn warnings conveyed higher urgency ratings than OneColumn warnings, a further study was conducted on layout types. If we continue to increase the number of columns of activated tactors to create tactile warnings, it is possible that participants rate the level of perceived urgency higher than TwoColumn warnings. This becomes the primary motivation of the second experiment.

To simplify the experimental design, the second experiment was identical in design and most procedures to the first experiment. The values of fundamental frequency and the duration of tactile stimuli in the first experiment were retained. For the three tactile parameters, all three activation levels were kept in the second experiment. However, the activation type parameter was restricted to FullBar only, as the results of the first experiment have revealed some drawbacks in the Target design. We summarize the design of three tactile parameters as follows: .

Activation levels

Three activation levels: Level 1, Level 2 and Level 3.

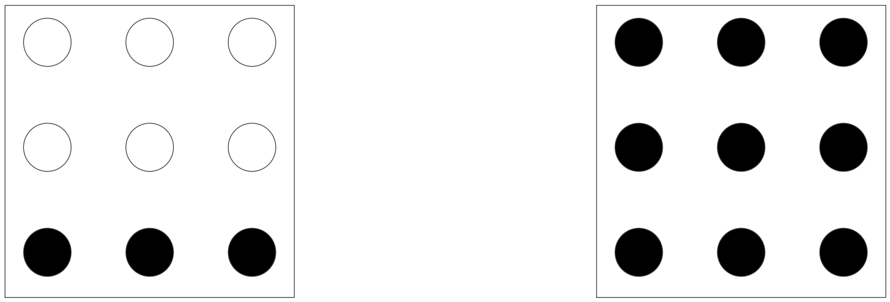
Activation types

The activation type was fixed as FullBar.

Layout types

The OneColumn and TwoColumn designs were retained. A third layout type was added in the experimental design. The ThreeColumn design was a combination of OneColumn and TwoColumn designs that tactors in three columns of the grid were activated simultaneously.

Figure 19 shows activated tactors of the grid when playing FullBar-ThreeColumn warnings at both Level 1 and Level 3.



FullBar-ThreeColumn-Level 1 FullBar-ThreeColumn-Level 3

Figure 19: Examples of the ThreeColumn layout type in Experiment 2.

4.3.2 Stimuli

As the activation type was fixed, the number of tactile parameters was reduced from three in the first experiment to two (layout type and activation level) in the second experiment. Thus, the term "design" and "layout type" were identical. For this reason, a "configuration" represented the combination of layout type $\times$ activation level. The summary of all 9 configurations in the second experiment was listed in Table 12.

Table 12: Configurations in Experiment 2.

Layout type (3)	Activation level (3)
OneColumn	Level 1
TwoColumn	Level 2
ThreeColumn	Level 3

4.3.3 Hypothesis

Our hypotheses in Experiment 2 are listed as follows:

1. More activated tactors should produce higher perceived urgency. This hypothesis should corroborate the finding of the first experiment. This suggests that ThreeColumn warnings should be perceived more urgent than TwoColumn warnings because more tactors are being activated.

2. Higher (closer to the shoulders) activation levels should convey higher perceived urgency. This hypothesis should corroborate the finding of the first experiment. Activation level should still be used to present three degrees of urgency for each design (layout type) similar to the first experiment.

4.3.4 Apparatus

All experimental equipment and setups in the first experiment was retained.

4.3.5 Procedure

A training section similar to the first experiment was included. In the familiarization task, participants were introduced with all the trials in the experiment. The results of the discriminability task in the first experiment showed that the design may need to be improved, as no significant result has been found. It was prudent to exclude the discriminability task from the design of the second experiment. Thus, participants started the urgency rating task immediately after the completion of the training section. After completing the urgency rating task, participants were asked to fill out a questionnaire on their overall preferences and annoyance levels between the designs. The questionnaire was similar to the one in the first experiment. An example of the questionnaire can be found in Appendix B.2

4.3.6 Participants

Ten students were recruited from the University of Waterloo to complete this experiment. Five participants were female and five participants were male. Table 13 summarizes the completion of each task. All participants chose the medium-size vest without adjusting the position of the tactor grid. All participants completed the urgency rating task and the questionnaire. Urgency rating data of all the participants were used in data analysis. Because one participant failed to provide adequate information on the questionnaire, only nine questionnaires were examined.

Table 13: Completion of Experiment 2.

Task	Number of participants
Completed urgency rating task	10
Completed the questionnaire	9

4.3.7 Results

Urgency Rating

The data was organized in a way similar to the first experiment. We calculated the mean rating of five replicates per trial. The mean rating was normalized using the approach in Appendix A.1. The Shapiro-Wilk test indicated that the distribution of normalized mean rating was not normal in the group of OneColumn $\times$ Level 1, $p = 0.006 < 0.05$. Distributions in the other groups were normal, $p > 0.05$. Either the OneColumn in layout types or Level 1 in activation levels can be excluded from data analysis as a correction. Because layout type was the most interesting variable in this experiment, Level 1 was removed from statistical model.

The statistical model was revised to three activation types (OneColumn, TwoColumn and ThreeColumn) $\times$ two activation levels (Level 2 and Level 3). Hence, the total number of configurations of tactile stimuli was changed from 9 to 6. A Friedman test showed that the participants rated all the trials consistently, Kendall's $W' = 0.833,\ x^2(11) = 41.646, p < 0.0005$. Data distributions in all trials were normal, $p > 0.05$. There was a homogeneity of variances, which was assessed by Levene's test for equality of error variances, $p = 0.999 > 0.05$.

A two-way ANOVA – activation type $\times$ activation level – indicated that there was a significant main effect for normalized mean rating in layout type, $F(2, 54) = 6.195, p = 0.004 < 0.05$ and activation level, $F(1, 54) = 6.551, p = 0.013 < 0.05$. The activation level by layout type interaction was not significant ($p = 1.183 > 0.05$), which means either of the two parameters has an independent effect on perceived urgency.

Figure 20 shows the clustered bar chart is provided for comparing the effects of activation level and layout type. Note that the Y axis represents only the Level 2 and Level 3 for activation levels. For comparison of the three layout types, A Turkey HSD post-hoc test showed that normalized mean rating was statistically significantly higher in TwoColumn warnings than OneColumn warnings, $M = 0.17, SD = 0.06, p = 0.026 < 0.05$. Normalized mean rating of ThreeColumn warnings was also significantly higher than OneColumn warnings, $M = 0.21. SD = 0.06, p = 0.005 < 0$. However, no statistical difference of the rating was found between TwoColumn warnings and ThreeColumn warnings.

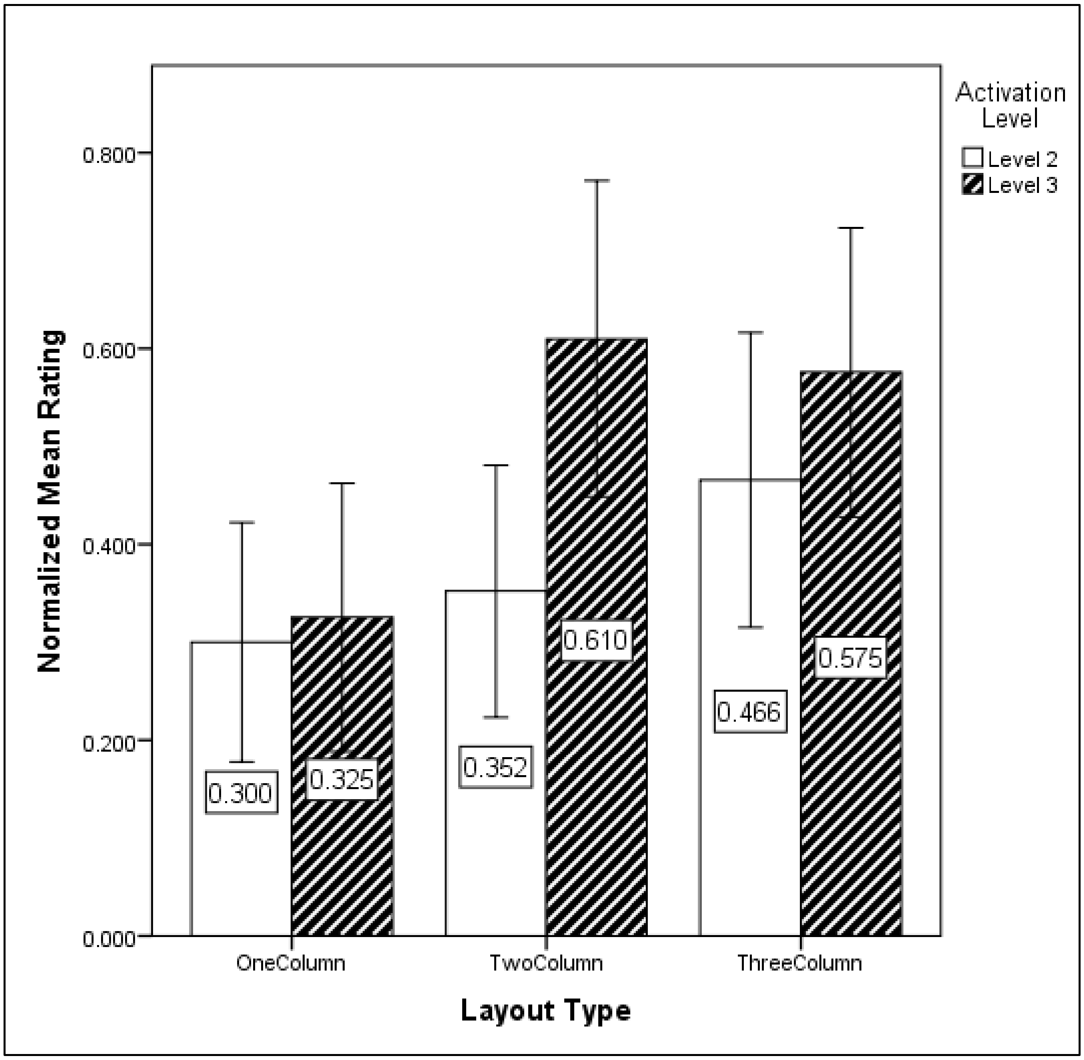


Figure 20: Clustered bar chart with confidence intervals of Experiment 2 (from SPSS).

Questionnaire - Preference

Participants were asked for their overall preference in the three designs (FullBar-OneColumn, FullBar-TwoColumn and FullBar-ThreeColumn) similar to the first experiment. A Pearson Chi-square test showed that there was no significant association between design and preference, $\chi^2 = 4.800, p = 0.091 > 0.05$. The descriptive statistical results are shown in Table 14.

Table 14: Preference for design in Experiment 2.

Design	Participants' Preference
FullBar-OneColumn	11%
FullBar-TwoColumn	22%
FullBar-ThreeColumn	67%

Questionnaire - Understandability

Participants were asked to choose the design which seemed easiest to understand. A Pearson Chi-square test showed that there was no significant association between design and perceived understandability, $\chi 2(1) = 4.000, p = 0.135 > 0.05$. Table 15 shows the perceived understandability of each design in the experiment.

Table 15: Understandability for design in Experiment 2.

Design	Participants' Preference
FullBar-OneColumn	11%
FullBar-TwoColumn	33%
FullBar-ThreeColumn	56%

Questionnaire - Annoyance

Each participant was asked to rank the perceived annoyance of the three designs using the same scale in Experiment 1 (Table 16). A Friedman test showed mean annoyance score was not significantly different in three designs, $\chi 2(2) = 0.222, p = 0.895 > 0.05$. No design was rated as "too annoying to use".

Table 16: Mean annoyance score of design in Experiment 2.

Design	M	SD
FullBar-OneColumn	2.111	0.928
FullBar-TwoColumn	2.000	0.707
FullBar-ThreeColumn	1.889	0.928

4.3.8 Discussion

Experiment 2 corroborates some findings in Experiment 1. First, the results of Friedman's test indicate that participants rated the tactile trials consistently. The results of the two-way ANOVA show that each of the activation level and layout type has an independent effect on perceived urgency of the tactile warnings.

The hypothesis that higher (closer to the shoulders) activation levels produced higher perceived urgency was partially confirmed, as Level 3 warnings were rated more urgent than

Level 2 warnings. However, as the degree of activation levels was reduced from three to two due to data correction, the effect of activation level was not fully revealed in this experiment.

Our hypothesis that more activated tactors produced higher perceived urgency was partially supported only, because TwoColumn warnings were rated more urgent than OneColumn warnings and ThreeColumn warnings were also rated more urgent than OneColumn warnings, but ThreeColumn warnings were not rated statistically more urgent than TwoColumn warnings (Table 17). This finding was surprisingly different from the comparison of the TwoColumn design versus OneColumn design in this experiment. A possible reason was that the middle column of tactors was not detected correctly in ThreeColumn warnings. Participants may not clearly distinguish the TwoColumn and ThreeColumn warnings. The same reason has been proposed in discussion of the effects of TwoColumn design versus OneColumn design in the first experiment. Unlikely, in the first experiment TwoColumn warnings were rated significantly higher than OneColumn warnings.

Thus the effect of the middle (spinal) column of tactors in either OneColumn warnings or ThreeColumn warnings was still remaining to be further explained. As commented in the questionnaire, two participants believed that the OneColumn warnings were “distracting”, and it was more similar to a “poke” rather than a “serious warning”. The anatomic review of the torso suggested that the spine was one of the most sensitive points to perceive tactile stimuli (Cholewiak, 2004). It was possible that a distracting activation in the middle column may reduce perceived urgency. In summary, further research should be carried out to see how vibration in the spinal column may affect perceived urgency of tactile stimuli. In this case however, designers should be very careful to present tactile warnings in the spinal column.

Table 17: Comparison of parameters contribute to the perceived urgency of Experiment 2.

	Comparisons (normalized mean rating, SD)
Layout type	ThreeColumn ($M = 0.520, SD = 0.211$) >
	TwoColumn ($M = 0.481, SD = 0.239$) >
	OneColumn ($M = 0.313, SD = 0.177$)
Activation level	Level 3 ($M = 0.503, SD = 0.239$) >
	Level 2 ($M = 0.372, SD = 0.195$)

4.4 Summary

In this chapter, the approach to design tactile warning based on perceived urgency was examined in two experiments. First we built the tactile display with a grid of tactors. The grid was mounted on a vest worn by participants in the two experiments. In the experiments, the urgency of tactile warnings was determined by the participants with a subjective rating scale. Results show that the proposed tactile parameters (activation level, activation type and layout type) have some effects on perceived urgency of tactile stimuli. Therefore, by manipulating these parameters we can encode perceived urgency information into tactile warnings. From the results of the first experiment, activation level can be encoded with three degrees of urgency, and several designs (the combinations of activation type and layout type) should be identified as different warning designs. Results of the analysis indicated that the FullBar-TwoColumn design performed solidly in all three activation levels. It is recommended at this stage. No significant effect on participant's response was found within all warning designs due to the lack of context in the experiment. Thus, the effectiveness of designed warnings in more complex environment and how perceived urgency matches the severity of risks and hazards require further investigations.

Chapter 5

Experiment 3: Urgency Mapping in the Context of UAV Flight Simulation

In Experiment 1 and 2, several tactile warning designs were developed and urgency information was encoded in the tactile stimuli. However, both experiments had a limitation that the designs were tested in the absence of a realistic context. There are two challenges to present tactile stimuli in this context. The first challenge is the encoded urgency information in tactile stimuli may not match the severity of risks that occurred in the context. The presentation of tactile stimuli required correctly mapping urgency with the severity of risks. The second challenge is the magnitude of hazards may change gradually. The activation duration of our designed tactile stimuli was short (180ms), so the tactile stimuli varied in temporal forms. To describe the changes of magnitude, the designed tactile stimuli should be presented in a sequence to generate tactile patterns (Jones, 2009). Therefore, a subjective rating task should be taken to examine if the tactile patterns are perceptually distinct in terms of urgency, and if the urgency of tactile patterns match the severity of risks that occurred in the context. If so, tactile patterns can be more suitable tactile warnings than designed tactile stimuli to alert the pilots with hazards and risks in this environment. No discriminability task was designed for Experiment 3.

In this chapter, we evaluate the designed tactile stimuli in a context of unmanned aerial vehicle (UAV) flight simulation. First, we introduce the design of tactile patterns using the tactile stimuli and urgency mapping with the hazardous events in the flight simulation. After that we discuss the apparatus, procedure and the results of the experiment.

5.1 Experiment 3

5.1.1 Design

The third experiment was conducted on an unmanned aerial vehicle (UAV) ground control station (GCS) simulator. Different from manned aircraft, the UAV is remotely controlled by pilots. It is important for the pilot to detect a deviation of the UAV as quickly and accurately as possible, especially when the UAV encounters extreme weather conditions. A tactile interface can be used as a viable alternative to present such deviation to the pilot, when the UAV visual interface was overloaded by a number of flight indicators.

This UAV GCS simulator was built on a professional software package named XPlane (www.x-plane.com). The XPlane emulated several typical courses of flight, such as take-off, cruise, diversion and landing. The basic flight scenario was fixed, which meant that the pilots were following the same directions in each scenario. During the courses of flight, the UAV may encounter extreme weather conditions, such as wind shears and turbulence. The term "critical event" referred to these weather conditions that resulted in the deviation of the UAV. The duration of each event was set as 10s. A total of four events differed in severity were programed in the system. Depending on the severity, these events may result in shudder, deflection or stall. Pilots should always monitor the attitude of UAV, especially when the events occurred. The attitude of the UAV may deviate from the desired path along three dimensions, namely roll, pitch and yaw.

Burns et al. (2011) suggested that tactile stimulation could be mapped to the aircraft attitude deviation. However one should consider the presence of tactile adaptation (Nafe & Wagoner, 1941) which may reduce human performance in response to unchanged tactile signals. As the attitude deviation was continuous, a necessary gap should be inserted between two tactile stimuli. Each of our tactile stimuli was designed with a 20ms silence before the tactile activation in stimulus. This allowed our tactile stimuli to be played consistently within the duration of events.

5.1.2 Urgency Mapping

Urgency mapping was conducted in three steps. First, we chose tactile stimuli from the previous experiments and mapped the urgent information to the magnitude of UAV attitude deviation. Secondly, we generated the tactile patterns by presenting the tactile stimuli in a sequence within the occurrence of each event.

The duration of each event was 10s. The most severe event was high turbulence, and the UAV would crash shortly after the event occurred. Low turbulence, high wind shear and low wind shear did not result in UAV crashes. Low turbulence and high wind shear were moderate severe. Low wind shear resulted in a mild deviation.

When we revisited the four events, we found that the magnitude of deviation varied during the occurrence of each event. To monitor the deviation, a plugin was installed on the simulator to record flight status data at a refresh rate of 200ms. The status data included air speed, heading, altitude and three dimensions of attitude deviation: roll, pitch and yaw. A calculation was developed to aggregate the three dimensions of attitude deviation to one variable, namely *Attitude Status*. The procedures to build the calculation can be found in Appendix C. Figure 21 shows the variance of *Attitude Status* in the high wind shear event. The y-axis shows the results of *Attitude Status*, and the x-axis shows the time (with a 200ms interval).

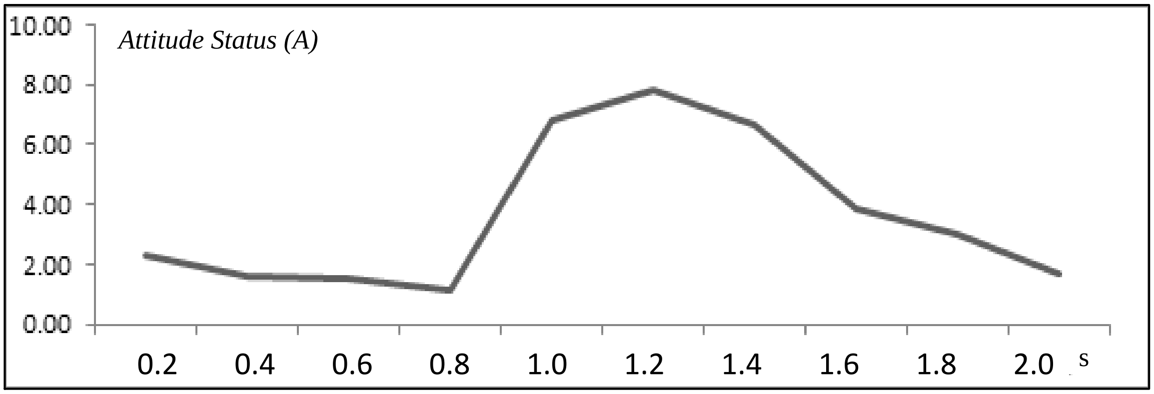
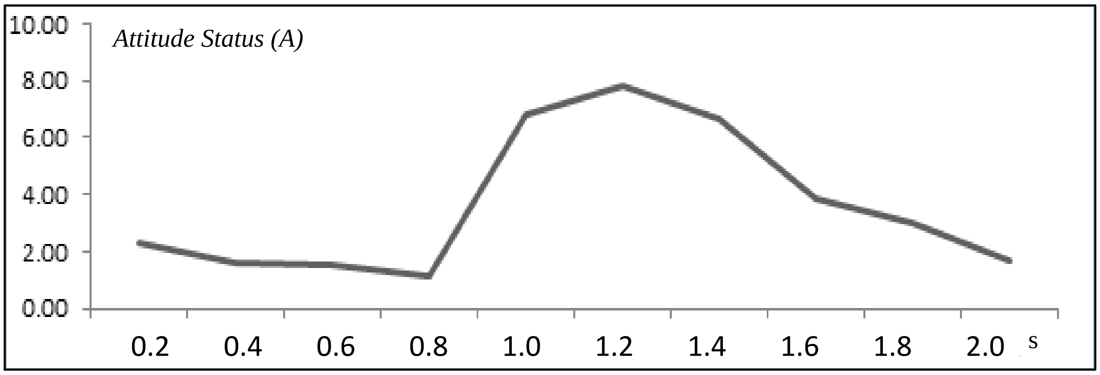


Figure 21: Example of *attitude status* in high wind shear.

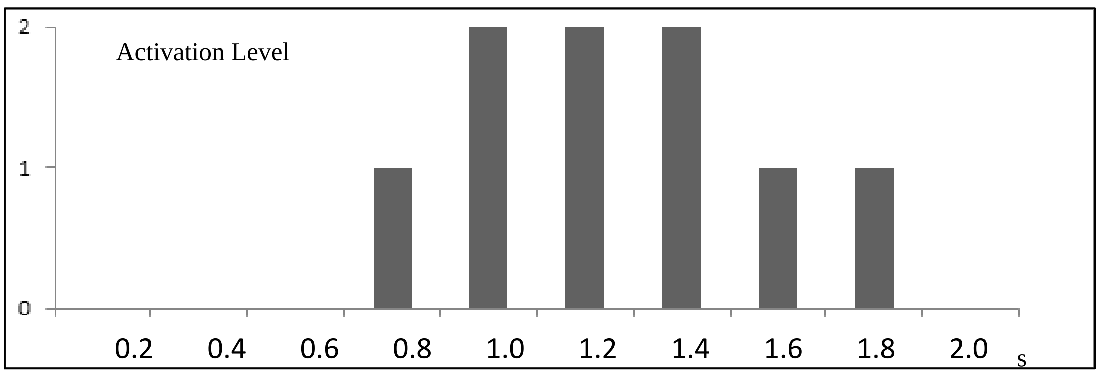
The presentation of tactile stimuli should describe the variance of *attitude status* in each event. The total duration of each tactile stimulus was 200ms, which was equal to the refresh rate of *attitude status*. The vibration strength was limited to avoid the presence of tactile adaptation (Nafe & Wagoner, 1941) by the 20ms silence before each time the tactors were vibrated. Thus, the tactile stimuli can be played consistently every 200ms.

Burns et al. (2011) suggested that the deviation of an aircraft from the correct direction can be represented by a column of tactors. As it has been recommended in Chapter 4, the FullBar-TwoColumn design was chosen. The temporal form of UAV attitude deviation was presented by designed tactile stimuli with different activation levels (Level 1, Level 2 and Level 3), and altogether a pattern of a variety of tactile stimuli was built for each event.

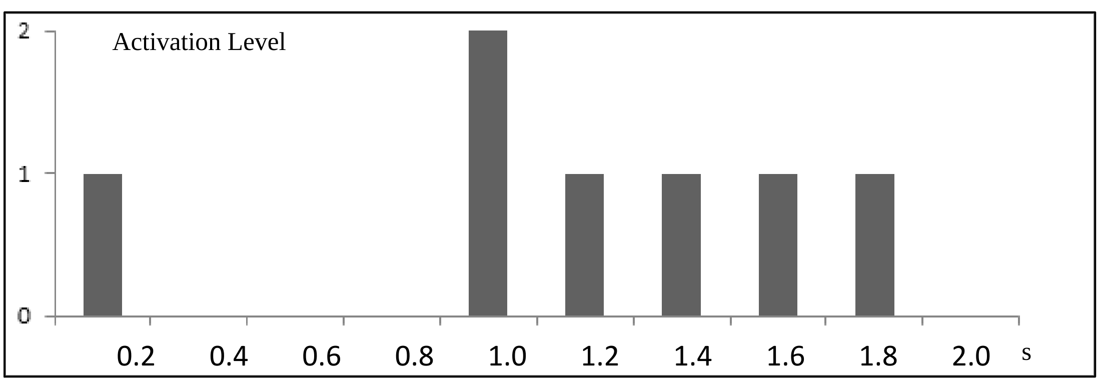
There were two approaches to generate a tactile pattern that present *attitude status* (Figure 22-a) by activation levels: in Figure 22-b, the pattern of tactile stimuli was fixed, which was specified by the severity of each event. The second option is described in Figure 22-c as a dynamic pattern, where each tactile stimulus in the pattern described the temporal form of *attitude status* at the moment. In this case, the temporal activation level of each tactile stimulus in the pattern was decided by the magnitude of current *attitude status*.



a) *Attitude Status*



b) Tactile Warning Presentation: Fixed Pattern



(c) Tactile Warning Presentation: Dynamic Pattern

Figure 22: Example of two ways of tactile warning presentation (tactile patterns).

One problem was found in the first approach for warning presentation: the pattern of *attitude status* was not only defined by weather conditions in the simulator. It was also affected by the current attitude of the UAV. For example, if the UAV's heading was changed when the event of low wind shear occurred, the form of *attitude status* can be slightly different.

The second approach to present tactile stimuli was adopted in our paradigm, as it provided a true description of the variance of *attitude status* in each event. There was one challenge to choose the second approach: the tactile patterns may not be perceptually distinct in terms of perceived urgency, as the tactile patterns of the same event in different flight scenarios may not be identical. For this reason, perceived urgency should be appropriately assigned with the severity of the event. To solve this, the magnitude of *attitude status* was first classified by three criteria, and they were associated with the three activation levels of tactile stimuli. The code of the XPlane was modified to allow the *attitude status* being calculated every 200ms (refresh rate). If the value of *attitude status* lay in a particular criterion, the tactile stimulus of the corresponding activation level was played. Table 18-a presents the proposed criteria for *attitude status*. The threshold for each criterion was carefully defined so that the tactile patterns for each event may be perceptually distinct in terms of perceived urgency. We expected activation levels of tactile stimuli within a low wind shear should be restricted at Level 1. Activation levels in either high wind shear or low turbulence should stay at Level 1 and Level 2. But we expected to see more Level 2 activations in low turbulence, as low turbulence was more severe than high wind shear. High turbulence should be the only event that can trigger activation Level 3. A comparison between severity of the event and the expected activation level (s) can be found in Table 18-b. The table also shows that if *attitude status* was lower than the first threshold (1.5), no tactile activation was presented (Figure 22-c). Figure 23 gives some examples of tactile patterns in four events. These examples confirmed the criteria we proposed.

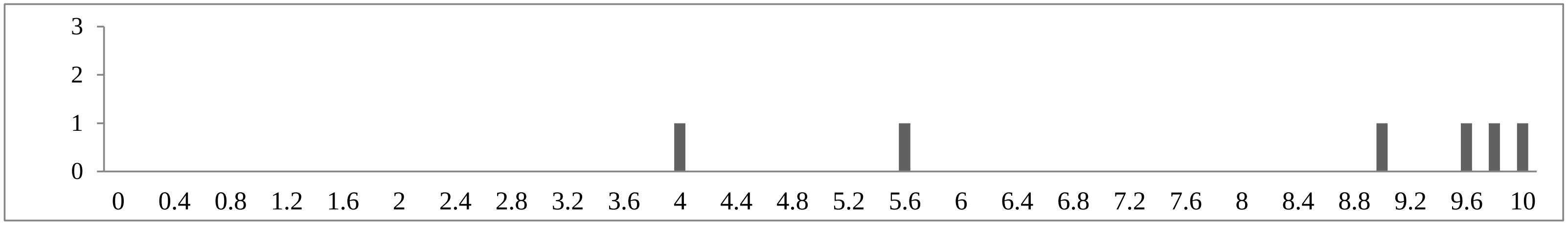
Table 18: Criteria of *attitude status* to match activation levels.

a) Proposed criteria of *attitude status*

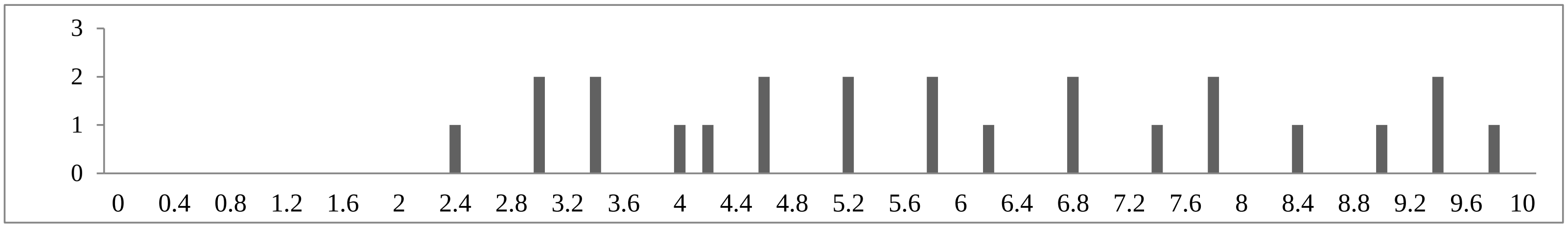
Attitude Status (A)	Activation Level
$1.5 \leq A < 3$	Level 1
$3 \leq A < 16$	Level 2
$A \geq 16$	Level 3

b) Expected activation level(s) in each event

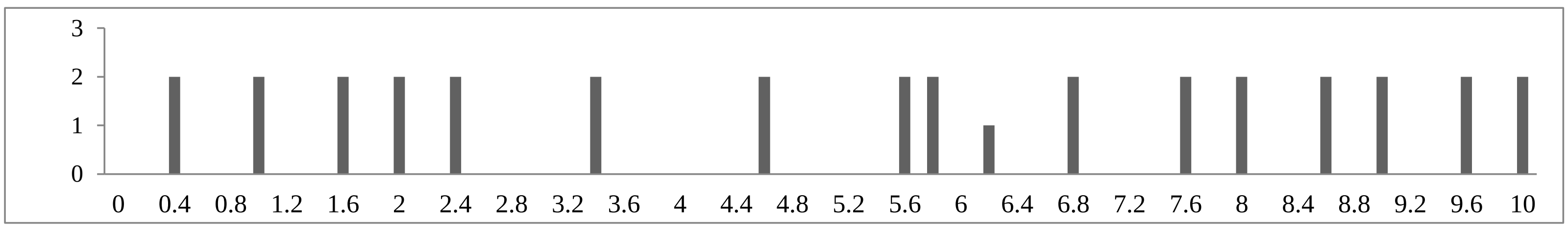
Severity of Event	Event	Expected activation level(s)
High	High turbulence	Level 1, 2 and 3
Medium high	Low turbulence	Level 1 and 2
Medium	High wind shear	Level 1 and 2
Low	Low wind shear	Level 1



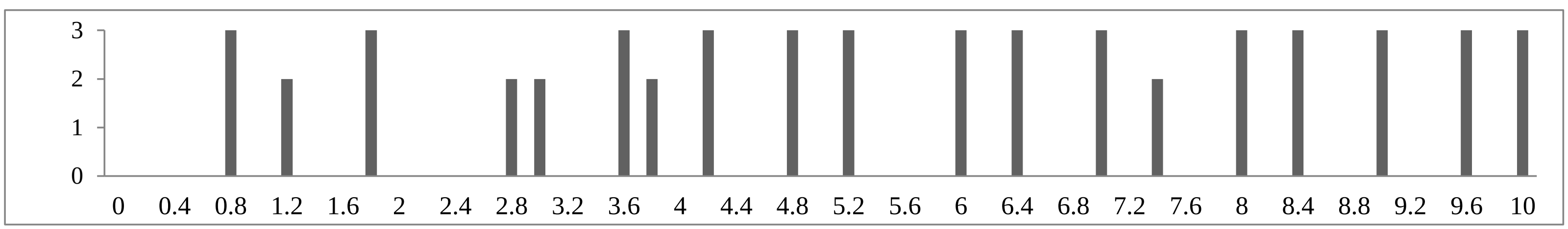
a) Low Wind Shear



b) High Wind Shear



c) Low Turbulence



d) High Turbulence

X-Axis: Second Y-Aix: Activation Level

Figure 23: Examples of tactile patterns in five events.

The results of the first two experiments revealed that the higher activation level conveyed higher perceived urgency. As the tactile pattern in each event was made by tactile stimuli at different activation levels, our hypotheses of urgency mapping for the severity of events was proposed as Table 19. Because the high turbulence event caused the UAV to crash, the urgency of this event should be perceived distinctly.

Table 19: Proposed urgency mapping to severity of events (hypotheses).

Severity of Event	Event	Proposed Perceived Urgency
High	High turbulence	High
Medium High	Low turbulence	Medium High
Medium	High wind shear	Medium
Low	Low wind shear	Low

5.1.3 Scenario

A special scenario was developed in the simulator to complete Experiment 3 (Figure 24). During the scenario, the monitors of the simulator were turned off to minimize the disruption of the visual information to perception of the tactile modality. First, the UAV was launched by the experimenter. When the UAV altitude reached 600ft, the UAV was tasked to the way point by the experimenter. The UAV made a smooth left turn and then flew to pass the way point. After

10s when the attitude became stable, the events were triggered in sequence. The duration of each event was 10s. A 15s idle time was added after the completion of each event to ensure the attitude returned to normal before the next event occurred.

However, one problem was observed in testing the scenario. The presentation of tactile stimuli was mapped to *attitude status*. When the UAV was making the left turn, the deviation happened and tactile activation was triggered. In real cases, the pilot might be able to identify the tactile activation as normal, at least different from other critical events. However, in this experiment participants were not presented with any visual information and the left turn was commanded by the experimenter. To simplify the experiment design, a temporary solution was taken that the left turn was regarded as a special event, but it was not be included in the data analysis. Participants were asked to rate the level of perceived urgency after the occurrence of each event. The same scale of 1 to 100 was used similar to Experiment 1 and 2. A complete list of all five events is presented in Table 20.

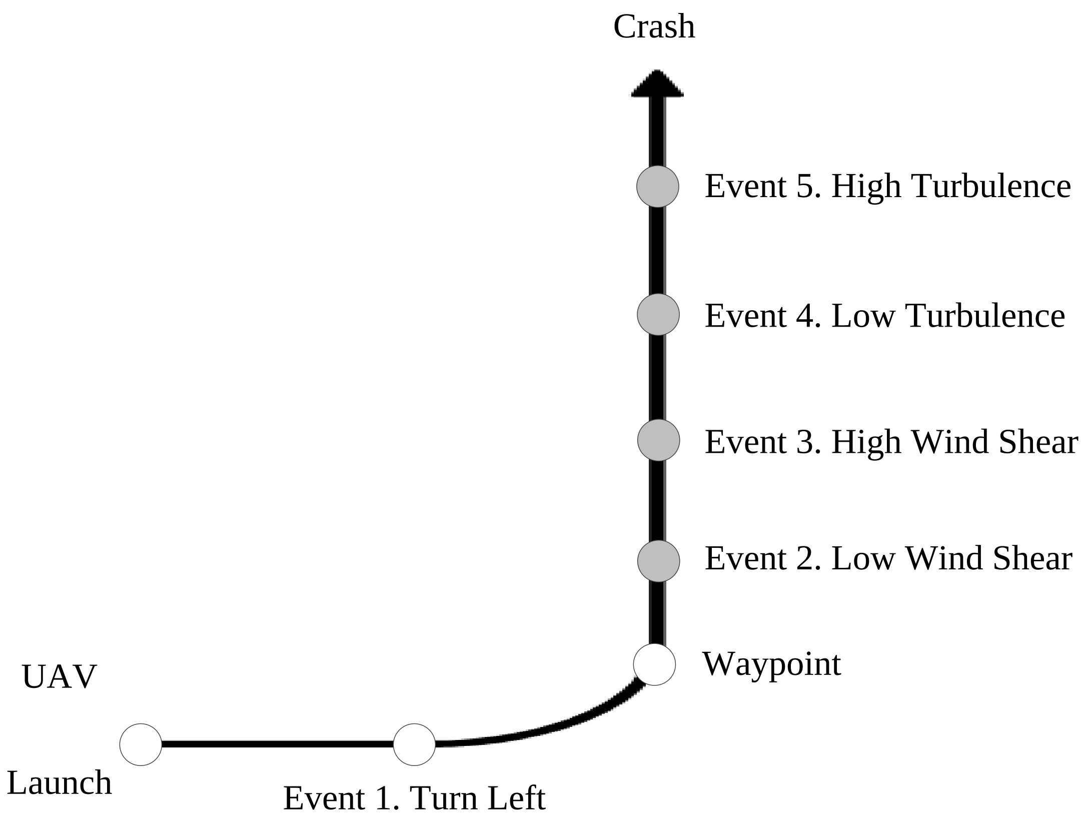


Figure 24: UAV scenario in Experiment 3.

Table 20: Events of the scenario in Experiment 3.

Event (Sorted by Occurrence)	Proposed Perceived Urgency
Event 1. Turn left	Not Applicable
Event 2. Low wind shear	Low
Event 3. High wind shear	Medium
Event 4. Low turbulence	Medium High
Event 5. High turbulence	High

5.1.4 Hypotheses

Our hypotheses in Experiment 3 are:

1. The urgency of each event should be perceptually distinct. High severity events should be perceived more urgent than low severity events, including medium and low severity events.
2. High turbulence causes the UAV crash, so it should produce significant higher perceived urgency than other events.

5.1.5 Apparatus

The experiment was taken after the completion of Experiment 2. The same group of participants participated in the third experiment. The set up of the experiment room was similar to that of Experiment 1 and 2. Following the second experiment, the GCS simulator was placed next to the experimental table. The Tactor Control Unit was connected to the GCS simulator by USB connection. The monitor was turned off during the experiment. The keyboard and mouse were not used (Figure 25).

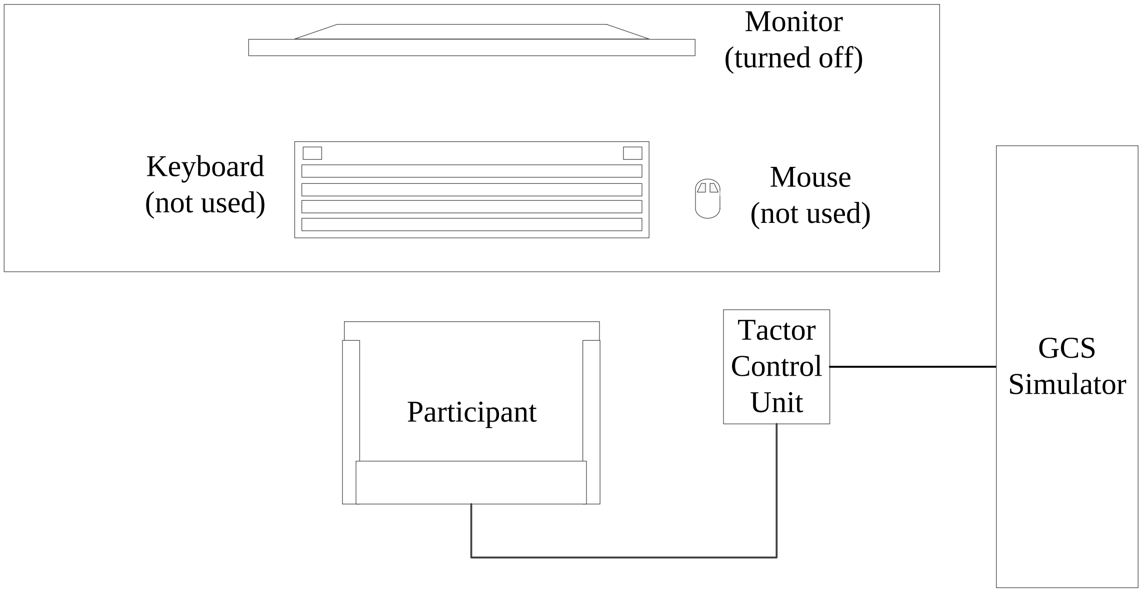


Figure 25: Experimental set up for Experiment 3.

5.1.6 Procedure

Before starting the experiment, each participant was given an answer sheet (Appendix B.3). In the answer sheet, participants were told that five events would be presented during the scenario, and the same scenario would be played twice. The experimenter started the simulation and launched the UAV on the GCS simulator. Participants were informed by the experimenter 15s before each of the events occurred. After the presentation of each event, the simulator was paused and participants were asked to rate the level of perceived urgency for the event. The same rating scale in Experiment 1 and 2, ranging from 1 to 100, was used. After the completion of urgency rating, the simulation was continued by the experimenter to prepare for the next event.

5.1.7 Participants

Due to time constraints, no training session was designed to familiarize the participants with the tactile stimuli. To share the training session of Experiment 2, the same group of participants in Experiment 2 participated in this experiment. All participants completed the urgency rating task for events (Table 21).

Table 21: Completion of Experiment 3.

Task	Number of participants
Completed urgency rating task	10

5.1.8 Results

Two corrections were made to the experimental data. First, the first event (turning left) was excluded from data analysis. We calculated the mean rating of two scenarios per event. After that the mean ratings were normalized with a calculation in Appendix A.1. The second correction was made when the rating of low turbulence was not normally distributed. Therefore, the data was corrected as transformed mean rating using the method in Appendix A.2. The results of Friedman test showed that the participants rated the level of perceived urgency consistently between the four events, Kendall's $W' = 0.818,\ x^2(3) = 24.526, p < .0005$.

The result shows that the distribution of transformed mean rating was normal (the *p*-value for each event was greater than 0.05). The test of homogeneity was passed ($p = 0.690$). A one-way ANOVA was run on transformed mean rating for four events. The results showed that there was a statistical significant difference of the transformed mean rating in events, $F(3,39) = 7.428, p = .001$.

A Tukey HSD post-hoc test showed there was a significant increase in transformed mean rating from the low wind shear event ($M = 0.243, SD = 0.196$) to the low turbulence event ($M = 0.625, SD = 0.239$), $p = .010$. There was a significant increase of transformed mean rating from the low wind shear event to the high turbulence event ($M = 0.744, SD = 0.331$), with a mean increase of 0.501, $p = .001$, and from the high wind shear event ($M = 0.428, SD = 0.240$) to the high turbulence event, $p = .041$. Other increases in transformed mean rating were found for each pair of low to high severe events, but none of them was statistically significant. Figure 26 shows the results of descriptive statistics.

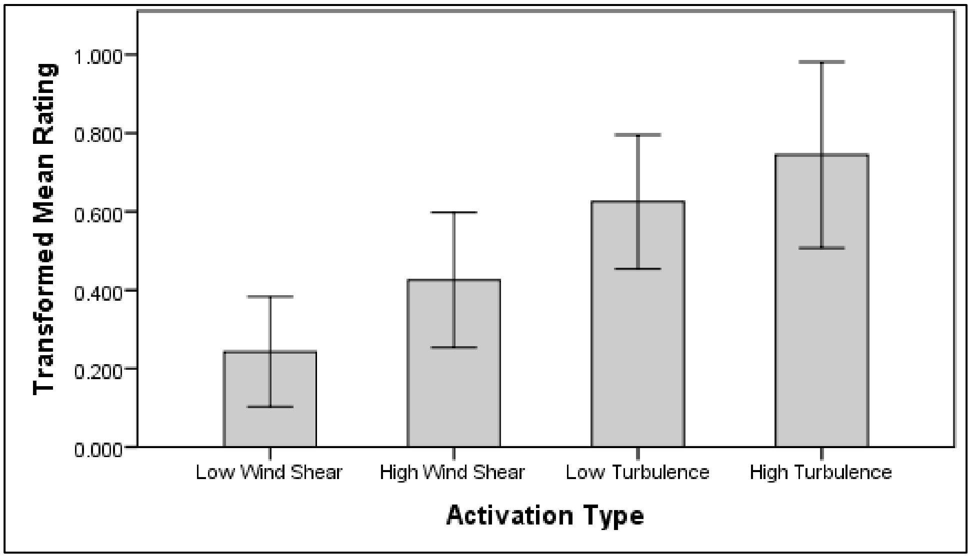


Figure 26: Transformed mean rating of events in Experiment 3.

5.1.9 Discussion

Generally, the results of Friedman's test revealed that perceived urgency rating for tactile pattern of each event was a viable approach. The urgency rating results showed that the high turbulence event produced significant higher urgency than the two wind shear events, which partially met the hypothesis. However, no significant difference was found between the high turbulence and low turbulence events. One possible reason for this was the tactile vibration in the turbulence events was strong, especially compared to the vibration in the low wind shear or high wind shear events. The density of tactile stimuli in the patterns was high. As only the FullBar stimuli were used to generate the tactile patterns, it was possible that the activation Level 1 and Level 2 were activated very often. A pressure adaptation may have occurred when constant vibrations were presented to the same location of the body (Nafe & Wagoner, 1941). It has been suggested by Van Erp et al. (2008) that if such vibration is presented for more than 200 ms, the pressure adaptation may occur. However, they also suggested that it is not necessarily to always move the tactors to another location on the skin, because the skin may also keep the sensitivity if some parameters of the stimuli are changed. The presumption has brought us back to the results of the first experiment that the FullBar was suggested as a more consistent design than the Target.

The finding suggests that further research is needed to generate a perceptually distinct tactile pattern. A past study (Brown et al., 2005) demonstrated that tactile patterns can be created by varying the vibration duration of tactile stimuli. Thus, to precisely create a tactile pattern which can be appropriately mapped with the severity of hazards, further experiments should be conducted on other tactile parameters, such as vibration duration. The results also suggest that the thresholds for *attitude status* should be improved, as the correlations of low wind shear versus high wind shear and low turbulence versus high turbulence were not clear.

One limitation we found was the lack of discriminability task. In the context of flight simulation, however, ideally the pilot should be presented with the visual displays as well, which may give the pilot some cues to detect different problems.

5.2 Summary

Experiment 3 examined the approach for mapping perceived urgency of designed tactile stimuli to the severity of hazardous events. The results showed that participants rated the four events consistently. This finding supported the current design to match tactile stimuli to the magnitude of UAV deviation. Some evidence was found that the high turbulence event was rated more urgent than the low and high wind shear events. This partially met the design requirement that the high turbulence event caused the UAV crashed, which was most hazardous in this environment. However, the design for tactile warning presentation within each event should be improved to give more comparisons between events in terms of perceived urgency. Further investigations may focus on revising the thresholds to trigger tactile stimuli on UAV deviation. For example, the lowest threshold was above zero. Thus, some mild deviation was ignored by the tactile interface.

Chapter 6

Conclusions

6.1 Summary

This thesis is an exploration into design tactile warnings that appropriately match the levels of urgency to the severity of the hazards. In Chapter 1 we presented two problems in the design of warnings as the presence of too much information can cause information overload. First, the visual interfaces, which were commonly used to present information, were overloaded frequently. The second problem was sometimes warnings did not match the severity of the hazard.

To solve these problems, we utilized an auditory warning design approach based on perceived urgency, and applied the principles to the design of tactile warnings. Our work started with a review of related works on auditory warning design based on perceived urgency and tactile interface design. The review showed that it is possible to establish a similar design approach in the tactile modality by adjusting warning parameters and urgency rating technique. The major procedures to design auditory stimuli based perceived urgency were retained, though there were a few revisions regarding modality-specified tactile parameters. The subjective rating technique, which quantified perceived urgency for tactile warnings, was implemented in the approach, and the same scale of rating in the auditory warning tactile was kept. Chapter 3 proposed a four-step approach to tactile warning design.

Three experiments were completed to support our approach. First, we developed a tactile display which was mounted on a vest. The tactile display was made of a grid of tactors. The vest provided a practicable display to present tactile warnings, and it was used throughout the three experiments. All three experiments (Chapter 4) were taken in a laboratory controlled environment. We developed a set of tactile warnings by manipulating three parameters of the tactile stimuli and examined perceived urgency of each stimilus with the subjective rating scale. The rating scale was inherited from the similar approach for auditory warning design with consistent results. The results revealed that perceived urgency was an important component in tactile warnings. By manipulating some parameters of tactile warnings, the level of perceived urgency changed regularly.

The most promising tactile warning design was examined in the third experiment (Chapter 5). In the third experiment, we examined the design of tactile stimuli in a context of UAV flight simulation. The tactile display was used to present UAV attitude deviation, and tactile warnings were developed to alert the pilot when the UAV encountered critical weather conditions (events) in the simulation scenarios. To present the variance of UAV attitude deviation, we quantified the deviation and generated tactile patterns by varying the level of urgency of several tactile stimuli. To verify the design, the third experiment was conducted to rate the urgency level of each tactile pattern when the events occurred. The same rating scale of the first two experiments was used. The results showed that some events were mapped correctly with designed tactile patterns, but further studies should be taken to improve the urgency mapping.

6.2 Contributions

From the results of the three experiments, the design of the tactile display seemed consistent. This supported the previous studies that tactile interfaces can be used to present meaningful information via the skin of human body. Thus, the thesis examined the potential of using the tactile modality as an alternative mode for presenting warnings, when there the information overload occurred on the visual interface.

The results of the experiments showed that the effects of tactile warnings were quantifiable. This supported the finding in the related works of the auditory warning design that perceived urgency was an important component in warnings, and the urgency information can be encoded in the warning stimuli at the design phase. The effects of some tactile parameters to perceived urgency of tactile stimuli were revealed. This finding was similar to the results in auditory warning design (for example: Edworthy et al., 1991). The major finding in the effect of tactile parameters was by increasing the number of activated tactors, higher urgency was conveyed. These parameters included activation type (FullBar warnings were rated more urgent than Target warnings) and layout type (TwoColumn warnings were rated more urgent than OneColumn warnings) in the first experiment. However, one exemption was observed in the second experiment, in that no statistically significant difference of urgency rating was found between ThreeColumn warnings and TwoColumn warnings.

The third experiment explored the way to map the level of urgency to the severity of hazard in a realistic context. The design of the experiment suggested that the presentation of tactile stimuli should be improved to present the varied magnitude of severity. The results of the third experiment showed that tactile warnings could be used to match some levels of severity, but future studies were needed.

6.3 Future Works

Future works building from this study are discussed as follows:

1. The parameters of tactile stimuli should be further explored. Any of the three parameters (activation level, activation type and layout type) describes the temporal form of the stimuli. Some parameters, such as fundamental frequency and vibration duration were fixed in

the current study, but they may also have some effects to the level of perceived urgency. Further studies should also include more complex parameters such as burst duration and waveforms as suggestion in Brown et al. work (2005).

2. The effect of the layout type is sill remaining to be further elucidated. The spinal column of tactors was activated in both OneColumn and ThreeColumn designs. However, the effects of OneColumn and ThreeColumn designs may be different. Designers should be cautious to use the spinal column to present tactile warnings.

3. The lack of content in the discriminability has called for further studies to explore the contribution of urgency information in human behaviors. The limitation of perceived urgency has been discussed in Chapter 2 that warning compliance can be a potentially interesting topic. Although an appropriate urgency mapping can be achieved, the effectiveness of warnings to behavioral compliance remained unknown. Some potential methods to improve the design of discriminability test can be just-noticeable difference (JND) and the magnitude estimation test which has been used already in auditory warning design studies (Edworthy et al., 1991).

4. The approach for urgency mapping should be improved. To ensure the tactile patterns are perceptually distinct in terms of perceived urgency, the design of the tactile patterns needs further improvement. The effects of the FullBar designs should be explored in further

experiments, because in the context of the third experiment, the use of FullBar designs may cause the risk of pressure adaption. Pressure adaption can be another interest of the further study, given us the opportunity to adjust the parameters of tactile stimuli and try to avoid such sensitivity decrease of the skin. Thus, designers should be careful to choose the FullBar designs to present the severity of hazard, especially in a scenario similar to Experiment 3.

Appendix A.1 Data Correction Methods

A.1 Normalization

The mean of ratings in the two scenarios has been used to give an overall rating. The mean value was then confined within 0 to 1 by the calculation below:

$$Normalized\ Mean\ Rating = \frac{Mean\ Rating - 1}{100 - 1}$$

Note that the rating for perceived urgency was set to range from 1 to 100. The value of "1" is defined as "the least urgent." Therefore the rating was normalized by subtracting the lowest limit and divided by the range. In this case the zero point of the rating was shifted to 0 from 1 and level of measurement became a ratio scale ranging from 0 to 1. The distribution for this type of data could be binomial and transformation was required to convey an analysis of variance.

A.2 Transformation

For the second correction the normalized mean value was corrected by an arcsine transformation. An angle transformation was conducted in analyzing data expressed as percentage and decimal:

Arcsine transformation:

$$\boldsymbol{Y = arcsin(X)}$$

$\boldsymbol{Y}$- Transformed rating

Appendix B Experimental Documents

B.1 Questionnaire of Experiment 1

Participant Number:____________

Overall, which design did you prefer? (Please circle one of the following)

Target, OneColumn   FullBar, OneColumn

Target, OneColumn   FullBar, TwoColumn

Which activation type did you prefer? (Please circle one of the following)

Target FullBar

Which layout type did you prefer? (Please circle one of the following)

OneColumn   TwoColumn

Any additional comments about your preferences?

Overall, which design did you find easier to understand? (Please circle one of the following)

Target, OneColumn   FullBar, OneColumn

Target, OneColumn   FullBar, TwoColumn

Which design did you find easier to understand? (Please circle one of the following)

Target FullBar

Which design did you find easier to understand? (Please circle one of the following)

OneColumn   TwoColumn

Any additional comments about your choices about understanding?

Please rank the designs by their annoyance level. (Please put a number next to each design for each category; 1 is the least annoying)

Target, OneColumn   Target, TwoColumn

FullBar, OneColumn FullBar, TwoColumn

Were any of the designs too annoying to use?

Any additional comments about your choices about annoyance?

B.2 Questionnaire of Experiment 2

Participant Number:____________

Overall, which design did you prefer? (Please circle one of the following)

FullBar, OneColumn

FullBar, TwoColumn

FullBar, ThreeColumn

Any additional comments about your preferences?

Overall, which design did you find easier to understand? (Please circle one of the following)

FullBar, OneColumn

FullBar, TwoColumn

FullBar, ThreeColumn

Any additional comments about your choices about understanding?

Please rank the designs by their annoyance level. (Please put a number next to each design for each category; 1 is the least annoying. Here "*Annoyance*" is considered as a NEGATIVE feeling.)

FullBar, OneColumn

FullBar, TwoColumn

FullBar, ThreeColumn

Were any of the designs too annoying to use?

Any additional comments about your choices about annoyance?

B.3 Answer Sheet of Experiment 3

Participant Number:____________

In the experiment the plane will encounter five events. The tactors on your vest will be activated during the events. Please fill out the urgency scale from how you feel right after each event. 1 means the least urgency, 100 means the highest urgency.

1st Event

Urgency [1-100] ____________________________

2nd Event

Urgency [1-100] ____________________________

3rd Event

Urgency [1-100] ____________________________

4th Event

Urgency [1-100] ____________________________

5th Event

Urgency [1-100] ____________________________

Were there any of the events the same?

No. Yes. (Please specify)

Appendix C

Calculation of Attitude Deviation

Two sets of data for roll, pitch and yaw were available in the UAV status monitor: the instantaneous value and the changing rate. We were able to describe the deviation by setting up a calculation with UAV status data. This calculation was used to show the magnitude of attitude upset and can be used as the reference to develop tactile warnings

From the results of the UAV status monitoring, we did not find an obvious change in the instantaneous value for roll, pitch, and yaw. The instantaneous value changes irregularly, even when the UAV is moving steadily and the value seems to be slow to react to abrupt attitude changes during events. In this case we chose the changing rate data as a reference. For another major correction, the rate of yaw changes infrequently within the events, so it has been removed from the final calculation for the same reason. The last correction is to use the absolute value of changing rates, as the data is presented with positive and negative numbers. From our paradigm our tactile display will not show the disorientation of deviation, so the rates of roll and pitch are aggregated in the calculation. The combination of UAV attitude status indicates the severity of deviation and the final calculation is presented as below:

$$\text{Attitude status (A)} = \frac{|rate\ of\ role| + |\text{rate of pitch}|}{2}$$

References

Arrabito, G. R. (2009). Effects of talker sex and voice style of verbal cockpit warnings on performance. *Human Factors, 51*(1), 3-20.

Arrabito, G. R., Mondor, T. A., & Kent, K. J. (2004). Judging the urgency of non-verbal auditory alarms: A case study. *Ergonomics*, 47(8), 821-840.

Bliss, J. P., & Dunn, M. C. (2000). Behavioural implications of alarm mistrust as a function of task workload. *Ergonomics, 43*(9), 1283-1300.

Brown, L. M., Brewster, S. A., & Purchase, H. C. (2005). A first investigation into the effectiveness of tactors. Paper presented at *the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. Glasgow, UK: IEEE.

Brown, L. M., Brewster, S. A., & Purchase, H. C. (2006). Multidimensional tactons for non-visual information presentation in mobile devices. Paper presented at *the Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. New York, USA:ACM.

Burns, C. M., Ho, G., & Arrabito, G. R. (2011). Mapping ecologically to modalities. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 55*(1), 335-339.

Burt, J. L., Bartolome, D. S., Burdette, D. W., & Comstock, J. R. (1995). A psychophysiological evaluation of the perceived urgency of auditory warning signals. *Ergonomics, 38*(11), 2327-2340.

Chapanis, A. (1994). Hazards assoicated with 3 signal words and 4 colors on warning signs. *Ergonomics, 37*(2), 265-275.

Chih-Yuan, H., Nikolic, M. I., & Sarter, N. B. (2001). Multimodal information presentation in support of timesharing and effective interruption management. Paper presented at *the Digital Avionics Systems, 2001*. Daytona Beach, FL: DASC.

Cholewiak, R. W., Brill, J. C., & Schwab, A. (2004). Vibrotactile localization on the abdomen: Effects of place and space. *Perception & Psychophysics, 66*(6), 970-987.

Edworthy, J. (1994). Urgency mapping in auditory warning signals. In *Human Factors in Alarm Design.* Taylor & Francis.

Edworthy, J., & Adams, A. J. (1996). *Warning design: A research prospective*. Taylor & Francis.

Edworthy, J., & Hellier, E. (2006). Alarms and human behaviour: Implications for medical alarms. *British Journal of Anaesthesia, 97*(1), 12-17.

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors, 33*(2), 205-231.

Gallace, A., Tan, H. Z., & Spence, C. (2007). The body surface as a communication system: The state of the art after 50 years. *Presence-Teleoperators and Virtual Environments, 16*(6), 655-676.

Gemperle, F., Ota, N., Siewiorek, D. (2001). Design of a wearable tactile display. In *Proceedings of the 5th IEEE International Symposium on Wearable Computers*. 5. Washington, DC: IEEE Computer Society.

Goble, A. K., Collins, A. A., Cholewiak, R. W. (1996) Vibrotactile Threshold in Young and Old Observers: The Effects of Spatial Summation and the Presence of a Rigid Surround. *Journal of the Acoustical Society of America, 99* (4), 2256-2269.

Guillaume, A., Pellieux, L., Chastres, V., & Drake, C. (2003). Judging the urgency of nonvocal auditory warning signals: Perceptual and cognitive processes. *Journal of Experimental Psychology-Applied*, *9*(3), 196-212.

Haas, E. C., & Casali, J. G. (1995). Perceived urgency of and response time to multi-tone and frequency-modulated warning signals in broadband noise. *Ergonomics, 38*(11), 2313-2326.

Hellier, E., Edworthy, J., Weedon, B., Walters, K., & Adams, A. (2002). The perceived urgency of speech warnings: Semantics versus acoustics. *Human Factors, 44*(1), 1-17.

Hellier, E.J., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors*, *35*(4), 693-706.

Ho, C., & Spence, C. (2007). Head orientation biases tactile localization. *Brain Research, 1144*, 136-141.

Hoggan, E., Brewster, S., & Acm. (2007). Designing Audio and Tactile Crossmodal Icons for Mobile Devices. In *Proceedings of the 9th international conference on Multimodal interfaces*. 162-169. New York, NY: ACM.

Ikeda, K. I. (1982). Warning of disaster and evacuation behavior in a Japanese chemical fire. *Journal of Hazardous Materials*, *7*(1), 51-62.

Jang, P. S. (2007). Designing acoustic and non-acoustic parameters of synthesized speech warnings to control perceived urgency. [Article]. *International Journal of Industrial Ergonomics, 37*(3), 213-223.

Jones, L. A., Kunkel, J., & Piateski, E. (2009). Vibrotactile pattern recognition on the arm and back. *Perception*, *38*(1), 52-68.

Jones, L. A., & Safter, N. B. (2008). Tactile displays: Guidance for their design and application. *Human Factors, 50*(1), 90-111.

Kaaresoja, T., & Linjama J. (2005). Perception of short tactile pulses generated by a vibration motor in a mobile phone. In *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. 471-472. Pisa, Italy: IEEE Computer Society.

Momtahan, K. L., & Tansley, B. W. (1989). *An ergonomic analysis of the auditory alarm signals in the operating room and recovery room*. Paper presented at the Presentation at the Annual Conference of the Canadian Acoustical Association.

Mondor, T. A., & Finley, G. A. (2003). The perceived urgency of auditory warning alarms used in the hospital operating room is inappropriate. *Canadian Journal of Anaesthesia-Journal Canadien D Anesthesie, 50*(3), 221-228.

Mori, Y., Tanaka, T., & Kaneko, S. (2011). *Design of vibration alert interface based on tactile adaptation model to vibration stimulation.* Paper presented at the In Proceeding of Symposium on Human Interface 2011, Orlando, FL.

Nafe, J. P., & Wagoner, K. S. (1941). The nature of pressure adaptation. *The Journal of General Psychology, 25*(2), 323-351.

Park, K. S., & Jang, P. S. (1999). Effects of synthesized voice warning parameters on perceived urgency. *International journal of occupational safety and ergonomics: JOSE, 5*(1), 73-95.

Patterson, R. (Ed.). (1982). *Guidelines for auditory warning systems on civil aircraft*. London: Civil Aviation Authority.

Perry, R. W., Greene, M. R., & Lindell, M. K. (1980). Enhancing evacuation warning compliance: suggestions for emergency planning. *Disasters, 4*(4), 433-449.

Rashid, R., & Wogalter, M. S. (1997). Effects of warning border color, width, and design on perceived effectiveness. *Advances in Occupational Ergonomics and Safety II*. Louisville, KY: IOS Press, and Ohmsha. 455-458.

Stuart, M., Turman, A. B., Shaw, J., Walsh, N., Nguyen, V. (2003). Effects of Aging on Vibration

Detection Thresholds at Various Body Regions. *BMC Geriatrics 3*(1), Retrieved January 28, 2013, from http://works.bepress.com/cgi/viewcontent.cgi?article=1003&context=bulent_turman.Suied, C., Susini, P., & McAdams, S. (2008). Evaluating warning sound urgency with reaction times. *Journal of Experimental Psychology-Applied, 14*(3), 201-212.

Summers, I. R., Cooper, P. G., Wright, P., Gratton, D. A., Milnes, P., & Brown, B. H. (1997). Information from time-varying vibrotactile stimuli. *Journal of the Acoustical Society of America, 102*(6), 3686-3696.

Van Erp, J. B. F., & Self, B. P. (2008). Tactile displays for orientation, navigation and communication in air, sea and land environments (No. TR-HFM-122). France: Neuilly-sur-Seine Cedex.

Van Erp, J. B. F., & Van Veen, H. A. H. C. (2003). *A multipurpose tactile vest for astronauts in the international space station.* Paper presented at the In Proceedings of Eurohaptics 2003.

Van Veen, H., & Van Erp, J. B. F. (2001). Tactile information presentation in the cockpit. In S. Brewster & R. MurraySmith (Eds.), *Haptic Human-Computer Interaction, Proceedings* (Vol. 2058, pp. 174-181).

Verrillo, R. T. (1966). Effect of spatial parameters on vibrotactile threshold. *Journal of Experimental Psychology, 71*(4), 570-&.

Wiese, E., & Lee, J. (2004). Auditory alerts for in-vehicle information systems: The effects of temporal conflict and sound parameters on driver attitudes and performance. *Ergonomics, 47*(9), 965-986.

Wu, J., Zhang, J., Yan, J., Liu, W., & Song, G. (2012). Design of a vibrotactile vest for contour perception. *International Journal of Advanced Robotic Systems*. Retrieved November 1, 2012, from http://cdn.intechopen.com/pdfs/40726/InTech-Design_of_a_vibrotactile_vest_for_contour_perception.pdf.