Interplay of Tactile and Visual Guidance Cues under Multimodal Workload

Mario Enriquez * ; Karon MacLean * ; Heather Neilson *

University of British Columbia

Abstract

Modern user interfaces – computerized, complex and time-critical – increasingly support users who multi-task; yet to do this well, we need a better understanding of how computer-user communication degrades with demand on user attention, and the benefits and risks of introducing new display modalities into high-demand environments, Touch can be a natural and intuitive locus of information exchange and is an obvious candidate for offloading visual and/or auditory channels. In this study we compared salience-calibrated tactile, visual and multimodal navigation cues during a driving-like task, and examined the effectiveness and intrusiveness of the navigation signals while varying cognitive workload and masking of task cues. We found that participants continued to utilize haptic navigation signals under high workload, but their usage of visual and reinforced multimodal navigation cues degraded; further, the reinforced cues under high cognitive workload disrupted the visual primary task. While multimodal cue reinforcement is generally considered a positive interface design practice, these results demonstrate a different view: dual-modality cues can cross a distraction threshold in high-workload environments and lead to overall performance degradation. Conversely, our results indicate that tactile signals can be a robust, intuitive and non-intrusive way to communicate information to a user performing a visual primary task.

Dept. of Computer Science, 201-2366 Main Mall, Vancouver, BC, V6T 1Z4, Canada {enriquez, maclean, haneilso}@cs.ubc.ca

1 Introduction

To accommodate the increased information exchange between humans and computers driven by technological advances, user interfaces are becoming more complex and users increasingly rely on them to perform tasks in parallel, and/or in distracting environments. While this communication can be improved by a well-designed interface, the increased complexity of interface or task context impacts usability. For example, in driver navigation aid systems it is critical that the information exchange remains reliable (signals are not missed) in many high-demand situations, but it must not interfere with safe driving. The haptic, or touch, sense may provide a solution for effective communication during driving, and in other high-cognitive-demand situations where both visual and auditory channels are overloaded. It is a natural and intuitive mode for us to gain information about the world; it is immediate and direct, and currently underutilized in modern interfaces. Haptic signals may thus be well-suited to communicate information reliably without causing unnecessary distraction.

Previous research has shown that haptic signals can be successfully used to communicate directional information, for example by giving cues to pilots to help them control the attitude of their aircrafts [1, 2]. Another study has found some differences between utilization of visual and haptic signals under workload in a driving task: hit rates in a detection task were more severely affected by difficult driving environments when the signals were visual compared rather than haptic [3]. This study was designed to determine whether peripheral haptic signals

could replace peripheral visual signals during a driving scenario; where the signals carried only presence/absence information.

Given that the primary tasks of both driving a car and piloting an aircraft currently rely predominantly on visual perception, we believe it may be useful for secondary displays such as navigation aids to communicate some information through haptic signals rather than the visual and audio information displays that are prevalent in modern systems. There is evidence that increasing drivers' cognitive workload in conjunction with a primarily visual task makes it more difficult for them to notice additional visual signals [4]; but it is not currently clear how signals in a *different* sensory modality will impact overall cognitive load.

In the experiment presented in this paper, we explored the premise that during a predominantly visual task, people's capacity to detect and respond to haptic navigation signals is less susceptible to the negative effects of cognitive workload than their ability to detect and respond to visual navigation signals. We also investigated the effectiveness of multimodal navigation signals, and the intrusiveness of all three types of signals. An intrusive signal is that which diverts attention from other important tasks. Here, a signal is considered intrusive if its presentation has a negative effect on the performance of an ongoing primary task of visual detection. This is a design trade-off, since clearly a more intrusive signal is more likely to be detectable under high workload.

We hope to demonstrate here that haptic signals are a robust channel for communicating navigational information during a primarily visual task, with implications for the design of navigational interfaces that are less intrusive yet more effective.

In the remainder of this paper, we will discuss related research; our setup, including methodology for simulating workload, perceptually calibrating the visual and haptic signals, and simulation of environmental masking; and our experiment design. Finally, we present and discuss our findings, their implications and our future work plan.

2 Related Work

Multiple resource theory suggests that when one modality is in use, it can be more efficient to communicate through a different modality [5, 6]. Alais, Morrone and Burr [7] found that each modality has its own attentional resources – there is less interference when two concurrent tasks involve separate modalities (vision & audition) rather than both involving the same modality, as long as the two tasks do not direct attention to different spatial locations. These findings suggest that while a person is engaged in an ongoing visual task, it may be more effective to communicate supplementary information via a different modality. In our experiment, an ongoing visual detection task is intended to simulate some aspects of the visual attention demands of driving; we investigated the effectiveness of communicating navigation information visually, haptically and through both modalities at the same time.

When two signals are presented at the same time, they may conflict with each other if they carry different types of information or require different responses, as suggested by the findings of Wickens and Hollands [6] and Alais, Morrone and Burr [7]. However, two signals presented in different modalities may also carry redundant information, in which case behavioural and neural imaging evidence suggest that cross-modal integration (additive or even multiplicative effects) may be expected. Gray and Tan [8] have successfully directed visual attention with clearly perceptible dynamic tactile cues in a low-workload setting, and Macaluso and Driver [9] have demonstrated that the intraparietal sulcus encodes spatial information for both visual and tactile stimuli – suggesting that similar perceptual processes might be involved in both

modalities. Stein and Meredith [10] trained cats to respond to visual and auditory stimuli, alone and in combination. They found that when the salience of the stimuli was low, presenting both stimuli simultaneously at the same spatial location increased response rates well beyond what would be predicted by combining the low response rates for the individual stimuli; whereas for higher salience signals, the improvement in response rates was not as dramatic.

Returning to the case of two competing signals (each related to a separate task), the change in communication efficiency that may occur when the two signals are presented through different modalities will likely only become apparent when resources are heavily used − that is, when cognitive workload is high. When cognitive resources are readily available, response to any combination of signals may be limited by physical perception ability alone. For the reminder of this paper, we will use Wicken's [5] definition of workload: the demand that is placed on mental resources.

Previous work has shown that increasing visual or audio cognitive workload makes it harder to notice visual signals. For example, Patten et al. [4] found that talking on a cell phone while driving slows reaction time for identification of visual targets, and that more complex conversations involving the driver have a larger effect than simpler conversations. Miura [11] found that in a real driving situation, increasing cognitive demand (by increasing the amount of traffic) increases reaction time to notice a small light projected onto the windshield, and decreases the maximum distance between the light and the gaze fixation point. Rantanen and Goldberg [12] have also shown that increasing cognitive workload reduces the size of the visual field.

Engstrom, Aberg and Johansson [3] investigated the effects of several forms of cognitive workload on visual and haptic signal detection performance (hit rate and reaction time) while participants drove (un-simulated) on the highway, on rural streets, or in the city. Hit rates for visual signal detection dropped significantly during city driving; for haptic signal detection, hit rates also dropped during city driving, but less severely. It should be noticed that hit rates for tactile and visual signals were not equalized for the baseline test. For both modalities, reaction times were higher in the city and when secondary workload tasks (counting backwards by sevens or dialling a cellular phone) were being performed.

Several studies suggest that haptic feedback is well suited to indicating directional information. Tan, Lim and Traylor [13] created a 3-by-3 tactor array (9 vibrotactile displays) which was used to indicate directions on a user's back, and Van Veen and van Erp [2] created a vest with 60 embedded tactors, which has been used to help helicopter pilots maintain stability in simulated flights. Rupert [1] showed that arrays of tactors on the torso can help pilots maintain an accurate sense of attitude (pitch and roll) in the absence of any visual attitude information.

In our experiment, we investigated the effects of cognitive workload on haptic and visual signals (as did Engstrom et al [3]); but as with more recent (unimodal and low-workload) directional studies, we used signals that carried extra information in the form of navigational cues. It is our belief that additional cognitive processing may be involved here than for a pure signal detection task. Further, we studied the effect of *multimodal* signals (both reinforced and alternated), and there found our most interesting results.

3 Approach

We wished to address two primary research questions. Are visual and haptic signal utility equally degraded by cognitive workload during a simultaneous visual search-respond task? Secondly, are visual and haptic signals equally disruptive of such a task?

3.1 Experiment Paradigm

We designed our experiment to directly compare accuracy of user navigational performance (measured by number of correct responses) in response to visual and haptic cues (alone or in various combinations) while manipulating the level of cognitive workload. To compare the intrusiveness of visual and haptic navigation signals on a visual search task, we used performance in a continuous workload task of detecting visual targets as an indicator of effective distraction level, where high distraction implies high intrusiveness of the signal.

We chose a spatial navigation paradigm – negotiating a computer-rendered maze, with turn-direction choices indicated only through navigation cues – because computer-aided navigation is a common task in which a variety of signals communicate information to human users, and navigation is often performed while drivers or pilots are under heavy cognitive workload from traffic and other factors (e.g. Baldwin & Coyne, 2003 [14]).

3.2 Tasks: Search with Secondary Navigation and Workload Manipulation

Because spatial navigation is carried out with primary visual attention focussed on the road or the instrument panel, we began with a foundational visual search and response task, with performance metrics that served as a measure of distraction-based degradation due to the secondary navigation task. In the navigation task, visual and/or haptic signals cued correct turning directions, and correct turns were thus considered a measure of the user's ability to perceive and respond to the signal under the conditions in effect. The factor of greatest interest here was the change in the signal utilization when cognitive workload was increased. Finally, we chose an auditory task to manipulate additional cognitive workload, in order to avoid biasing either the visual detection task or the haptic and/or visual navigation task.

3.3 Simulating Environmental Noise in Navigation Cues

In real scenarios (for example, everyday driving), visual and haptic cues would be perceived with a background of sensory environmental noise: a driver processes a continuous, distracting visual stream while feeling road vibration on the steering wheel. We therefore strove to simulate a realistic level of environmental masking, aiming for signals that were noticeable most but not all of the time.

Pilot studies showed large individual differences in visual and haptic stimulus detection ability. In order to make an objective comparison of the effects of workload on the two navigation signal types, we devised and carried out a custom calibration procedure to adjust stimulus-to-noise ratios so that each participant would have the same baseline performance level.

3.4 Factor Manipulations

We used a 4x2 factor design, with four multimodal variants in navigation signal and two levels of cognitive workload (see Table 1). In conditions employing the haptic-only signal type (denoted by 'H'), all of the navigation signals in the trial block were haptic, and in conditions employing the visual-only signal type ('V') all of the navigation signals were visual. We also included a reinforced signal type ('H+V') to investigate the effects of cross-modal integration; both a haptic signal and a visual signal were presented simultaneously for every trial in the block. Finally, we included a mixed signal type ('H|V') where either a haptic signal or a visual signal was presented for every trial, to explore the impact of broadening attentional requirements. Additional cognitive workload was either present or absent (denoted by the subscripts WL+ and WL- respectively).

Condition		Description
Haptic-Only (H)	Workload (WL+)	Vibration on left or right finger indicates turning direction
	No Workload (WL-)	
Visual-Only (V)	Workload	Triangle on screen indicates turning direction
	No Workload	
Reinforced (H+V)	Workload	Vibration and triangle both indicate turning direction
	No Workload	
Mixed (H|V)	Workload	Either a vibration OR a triangle indicates turning direction
	No Workload	

Table 1: The eight experimental conditions.
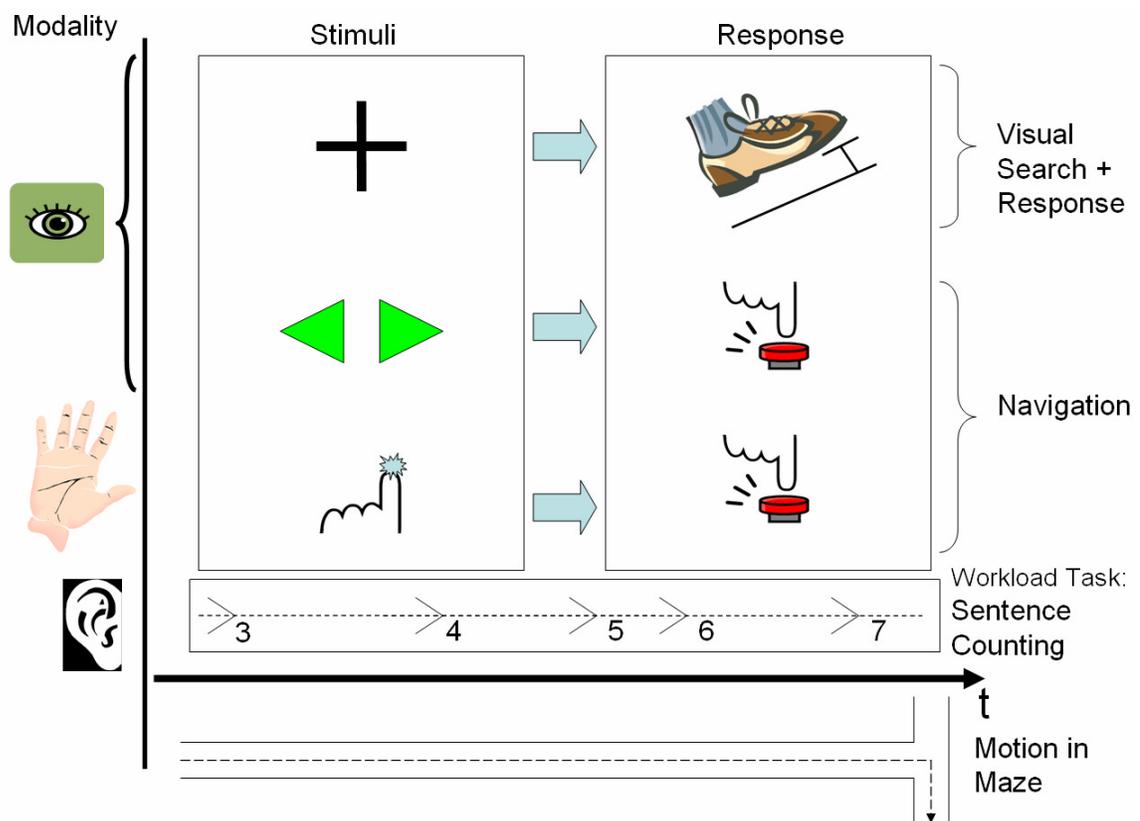
Figure 1: Schematic representation of one trial. The visual primary task was to notice the presence of a small crosshair embedded on the maze walls and react to it by pressing on a pedal. As the participant also listened continuously to a recording and counted the number of sentences (workload task, present in some conditions), he was presented with a tactile (buzz on either the left or right index finger) or visual (green directed triangle) navigation cue indicating the direction to turn at the next intersection. Position of the icons in the figure along the time axis indicates relative times at which these events might occur; in experiments the visual primary task (crosshair icon) occurred at random intervals, and the user chose when to respond to the directional cues pressing the corresponding turn button.

4 Methods & Materials

The visual search and response, navigation and workload tasks are depicted in Figure 1 and described in detail in following sections. For the ongoing, primary visual search task (in all conditions), participants were instructed to watch for targets (crosshairs) on the maze walls and respond by immediately stepping on a pedal. For the navigation task, we presented a tactile, visual, or both signals as participants approached each intersection in the maze. Participants responded by pressing left or right buttons to turn in the respective direction at the intersection. In all conditions, participants listened to spoken passages through a set of headphones. In trials with the added cognitive workload condition, they were asked to count the number of sentences they heard and report the number at the end of the trial block.

4.1 Apparatus

Physical setup

Participants were seated at a table with each hand resting comfortably on its own haptic display box (Figure 3), with the index finger on the vibrotactile display and the middle finger on the button. The two boxes were fixed to the table 28 cm apart. Participants viewed the maze on a 17" LCD monitor approximately 60 cm away from the participant, and listened to recordings of spoken passages for the workload task through a set of noise-cancelling headphones.

Two types of vibrotactile displays were used for the experiment. The target tactile signals were displayed using Audiological Engineering model VBW32 skin transducers [15] driven through a SoundBlaster Audigy 2 sound card. These voice-coil-based transducers are capable of producing precisely timed (on/off within 2 ms) waveforms at a useful range of frequencies and amplitudes, with maximum efficiency at 250 Hz. Tactile noise was generated by using T.P.C. model FM37E flat coreless vibration motors. These pager motors oscillate at approximately 133Hz at a fixed amplitude.

Maze

Participants advanced through a virtual maze rendered three-dimensionally, with a first-person point of view (Figure 2). Every intersection was a "T" where only left or right turns were possible – if the participant reached the wall and did not turn, he stopped. The maze intersections were generated randomly in real time, so in general the maze was not physically realizable. Participants were not allowed to back up and re-approach a specific intersection.
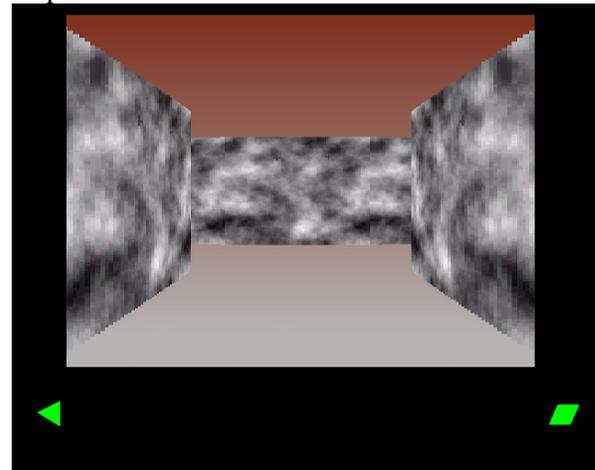
Figure 2: Screen shot of the maze, with a visual left-turn signal displayed. The turn signals were presented on their respective side of the screen while a different shape of similar salience was presented simultaneously on the opposite side of the screen to avoid the visual signal being identified by its location.

Responses

The experiment required participants to turn left or right by pressing a button on the haptic display box with the middle finger of the left or right hand respectively; they responded to the primary visual detection task by activating a foot pedal with one foot (chosen by participant) upon sighting a crosshair, and typed answers to a series of subjective questions at the end of each block

of trials using a standard keyboard.
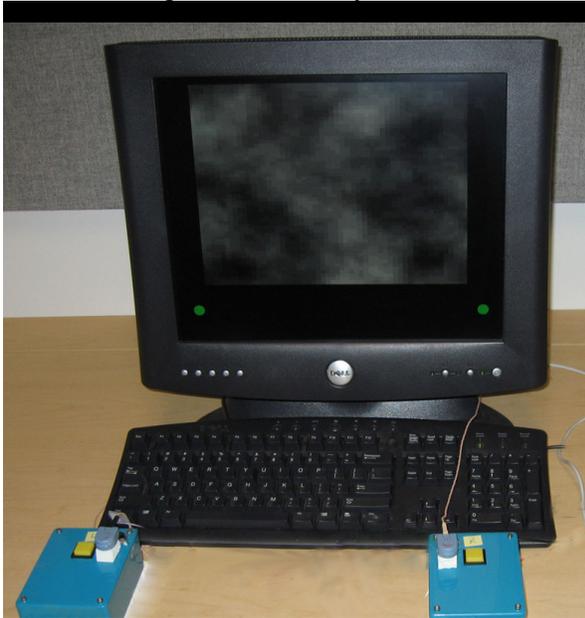
Figure 3: Participants placed one hand on each haptic display box and viewed the maze on the computer screen.

Responses

The experiment required participants to turn left or right by pressing a button on the haptic display box with the middle finger of the left or right hand respectively; they responded to the primary visual detection task by activating a foot pedal with one foot (chosen by participant) upon sighting a crosshair, and typed answers to a series of subjective questions at the end of each block of trials using a standard keyboard.

4.2 Visual Detection Task

Small crosshairs of a colour similar to the maze walls appeared randomly on the maze wall at the end of the corridor in 40% of the trials (Figure 6Figure 1). Participants were asked to watch for these targets and respond by pressing a foot pedal when they saw one. Two targets were always presented simultaneously in mirror locations on the left and right walls, to prevent the turning direction from being influenced by target asymmetry. Pairs of targets could appear in one out of four possible locations.

4.3 Navigation Task

Haptic and/or visual navigation signals indicated the direction to turn at each intersection in the maze. A cue was presented in one or both modalities before each maze intersection; the timing of the signal was randomized during an interval of 0.5 and 4 seconds prior to reaching the next intersection.

Visual signals consisted of triangles appearing in the left and right corners of the screen below the maze: a triangle pointing to the left in the left location indicated an upcoming left turn, and similarly for a right turn. To make the visual signals noisy, we also displayed a variety of randomly shaped but similar-salience distractor images in these locations in rapid succession (Figures 2 and 5), using the Rapid Serial Visual Presentation technique [16]. This limited the amount of time available to perceive and process each image and distinguish navigation cues from noise. By decreasing the presentation time for each visual image, we increased visual noise; making the targets harder to detect.

A haptic signal (target) was a 200ms, 250Hz vibration presented to the index finger of the participant's left or right hand using the high performance voice coil tactile display. A vibration on the left index finger indicated a left turn at the next intersection, and a vibration on the right index finger indicated a right turn. To make the haptic signals noisy, we applied a uniform level of background vibration using the pager motors oscillating at 133Hz, affixed to the outside of each haptic display box (Figure 5). Haptic signal-to-noise ratio was varied by adjusting the target haptic signal amplitude (presented through the voice-coil tactile display).

Signal Salience Calibration

Stimulus-to noise ratios for the navigation cues were adjusted independently for each participant such that by the end of the calibration phase and without the workload task, a participant would respond correctly to

visual navigation signals delivered alone approximately 80% of the time, and likewise for haptic navigation signals delivered alone. Effects of adding the workload task could then be directly compared in terms of resulting performance in both the primary visual detection and navigation tasks.

This calibration (in both modalities) was accomplished using an adaptive method (Parameter Estimation by Sequential Testing, or PEST, developed by Taylor & Creelman [17]) where signals are adjusted successively during a number of trials until the desired level of performance is reached. Following application of this procedure, we set the amplitude of the haptic signals to a median of 0.5 Volts, varying from 0.2 to 2, and the duration of presentation for the visual shapes to a median image display interval 155 ms, from 68 to 563 ms.
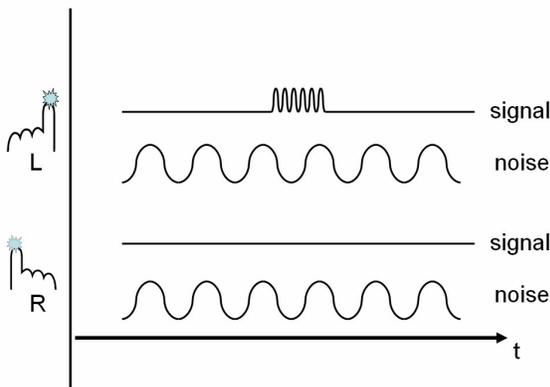
Figure 4: Haptic noise (ongoing vibration from pager motor) and a haptic left-turn signal (occasional burst signal from voice coil tactor). Relative signal amplitudes were adjusted for each participant but are approximate and representative.

4.4 Workload Task

In all conditions, participants listened to spoken passages through a set of headphones. In the added cognitive workload conditions, they were asked to count the number of sentences they heard.
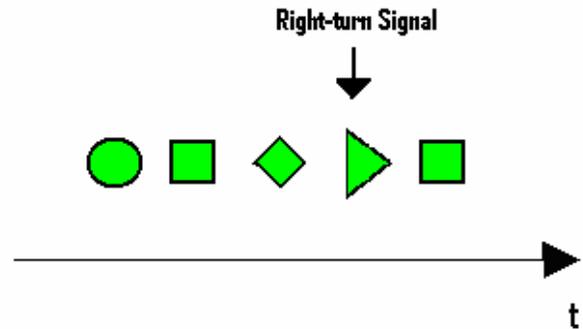
Figure 5: Rapid serial visual presentation, including a right-turn signal.

This task required participants to maintain a count in memory throughout each trial block, thus creating a continuous workload. At the end of the trial set, participants reported that number and indicated whether or not they had lost count. We did not consider the difference between the actual number of sentences heard and the number counted to be an indication of performance on the workload task: losing count of the sentences would likely result in guessing; in which case the number reported would not be a meaningful measure of performance. The spoken passages were taken from news articles, and read / recorded by the experimenter. The content of the passages chosen was mundane, and topics likely to arouse emotion or substantial interest were avoided. The average sentence length for each passage was between 19 and 22 words (median = 21.3), and no sentences contained less than 10 or more than 35 words.

4.5 Procedures

A trial consisted of a user experiencing one navigational intersection, and blocks in most cases consisted of 30 such trials. Post-training, a 30-trial block typically took about 90 seconds to execute.

For training, participants first listened to instructions from an audio recording. Participants were instructed to place an emphasis on visual target detection

(crosshairs) specifically. They were told that all other tasks were secondary and equally important amongst themselves. Participant instructions were carefully designed and presented via a recording to avoid confusion and/or experimenter influence on task emphasis [18].

Participants then practiced navigating the maze. For the first two 30-trial practice blocks (one with visual navigation cues (V) and one with haptic cues (H), order counterbalanced), there was no spoken passage playing, no background vibration (noise) on the haptic navigation cue display boxes, and the shapes on the screen changed slowly. Participants then practiced listening to a passage and counting the number of sentences they heard (Workload Task) with no navigation or visual task. Finally, two inclusive practice blocks were done – V and H – with the background vibrations on, the shapes changing quickly, and counting sentences of a passage playing through the headphones. Thus, one typical training sequence would be

V $\rightarrow$ H $\rightarrow$ Workload task $\rightarrow$ V_{WL+} with noise $\rightarrow$ H_{WL+} with noise,

but for half of the participants, the order of the V and H training blocks was reversed.

The next phase of the experiment was calibration; the haptic and visual signals were calibrated separately, so that the participant was able to navigate 80% of the turns correctly. Participants listened to spoken passages during the calibration but were not required to count the sentences in them, as was the case in non-workload experimental blocks. This was done to ensure that any differences observed with the introduction of the sentence counting task were due to the added cognitive workload and not to auditory noise or other perceptual factors.

Finally, eight blocks of experimental trials (Table 1) were carried out in random order for each participant, with a short break

(~0.5 minute) between each. The H, V and H+V blocks had 30 trials (30 intersections) each. The H|V blocks consisted of 30 haptic trials and 30 visual trials, for a total of 60 trials. In all cases half of the signals indicated left turns and half indicated right turns.

At the start of each block, a pop-up message box told participants what type of signals they would receive in order to minimize the adaptation period at the beginning of each block. Following each set of trials, a pop-up message box asked participants to estimate (as a percentage) how often they knew which way to turn, and in the cognitive workload conditions, asked them how many sentences they heard and whether they had lost count.

Once at the end of the experiment, participants were asked to fill out a questionnaire containing multiple choice questions about the H and V navigation signals, with respect to how comfortable they were and the difficulty of attending to them, and open-ended questions where they described the strategies they used.

4.6 Measures

For each experimental condition, performance in the navigation task was measured as the % of correct turns in a given block. Reaction time is not a meaningful measure for the navigation task, since participants were asked not to turn after receiving a signal until they reached the next intersection. For the visual target detection task, the percentage of visual targets detected out of all targets presented during the trial set was measured, and the mean time between the appearance of the target and the response was computed. Confidence was indicated by the participant's report (as a percentage value) of how often he or she knew which way to turn for each condition.
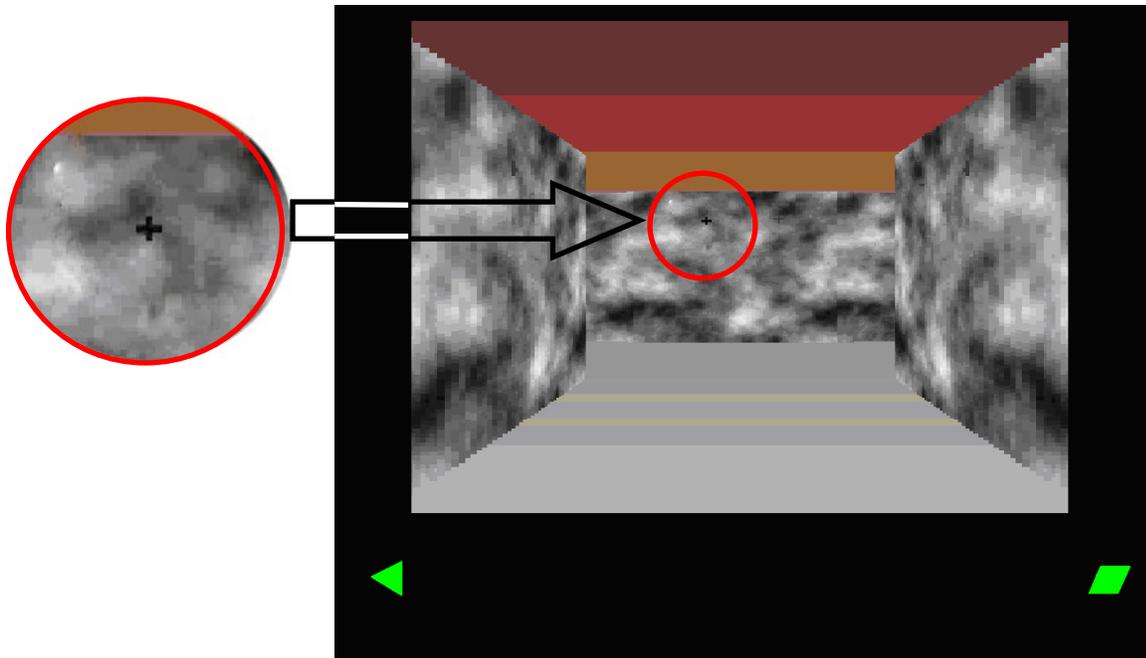
Figure 6: Close-up view of the crosshairs for the visual target detection task

SPSS	Condition	Mean Correct Turns (%)	Std. Dev.	Mean Visual Search Hits (%)	Std. Dev	Mean Pedal Reaction Time (s)	Std. Dev	Mean Confidence (%)	Std. Dev
w1_c1	H wl-	81.27	17.82	59.94	33.02	1.68	0.18	69.33	19.72
w1_c2	V wl-	82.57	9.04	57.55	35.52	1.66	0.37	56.79	20.98
w1_c3	H+V wl-	90.50	11.77	62.65	29.23	1.68	0.30	68.75	17.60
w1_c4	H|V wl- (H)	73.60	17.45	56.54	32.83	2.11	1.09	49.75	15.84
w1_c5	H|V wl- (V)	71.03	11.82						
w2_c1	H wl+	79.73	12.21	60.18	32.78	1.60	0.36	62.08	19.01
w2_c2	V wl+	74.10	11.15	58.97	31.54	1.53	0.52	48.83	22.64
w2_c3	H+V wl+	78.47	11.99	43.32	36.33	1.72	0.30	59.67	15.62
w2_c4	H|V wl+ (H)	72.30	19.45	58.79	36.21	1.55	0.36	45.00	22.64
w2_c5	H|V wl+ (V)	73.83	7.68						

Table 2: Means and standard deviations for all conditions and all metrics.

Metric	Condition 1 (best performance)	Condition 2 (lower performance)	p-value
Correct Turns	wl-	wl+	0.028
	Type		0.039
	H+V	H|V (V)	0.034
	H|V wl- (V)	V wl-	0.018
	H|V wl- (V)	H+V wl-	0.041
	V wl-	V wl+	0.026
	H+V wl-	H+V wl+	0.022
Visual Search	workload/type interaction		0.015
	H+V wl-	H+V wl+	0.007
Confidence	wl-	wl+	0.008
	type		0.008
	H	H|V	0.005
	H+V	H|V	0.007
	H wl-	H|V wl-	0.006
	H+V wl-	H|V wl-	0.004
	H+V wl-	H+V wl+	0.013

Table 3: List of significant pairwise differences between conditions, with p-values. The better-performing condition is always listed first.

5 Results

Thirteen individuals (seven male, aged 18-75 years with median 28.2, mostly university students) participated in this study. Most had no experience with tactile displays, and the remainder had moderate experience. Eight played video games less than once per month, and the remainder far more than this. Each participant was paid $10 for a 1-hour session.

Experiment results for all metrics and all conditions are listed in Table 2; significant post-hoc comparisons are detailed in Table 3. These results are elaborated by task in the following sections.

5.1 Visual Detection Task

In the primary task of visual target detection, average performance was 59.2% for conditions without the workload task, and 55.3% for those which included it. However, this difference was due mostly to the difference found in the reinforced signal condition (62.7% without workload vs. 43.3% with workload). These results are further discussed in section *6.3*.

A repeated-measures ANOVA on percentage of visual targets correctly identified showed a significant interaction between presence of the workload task and navigation signal type (F = 6.041, p=0.015, partial η^2 = 0.668). Target detection performance is shown in Figure 7; for the mixed condition (H\|V), trials with haptic signals are not separated from those using visual signals because the detection hit rate was computed for the whole duration of each trial block. Multivariate tests indicated that workload had a significant effect on identification performance in the reinforced condition H+V ($F_{1, 11}$ = 10.9, p=0.007), where performance decreased with workload; but that it was not a significant factor for the other types of signals.

A two-factor repeated-measures ANOVA found no significant differences across signal
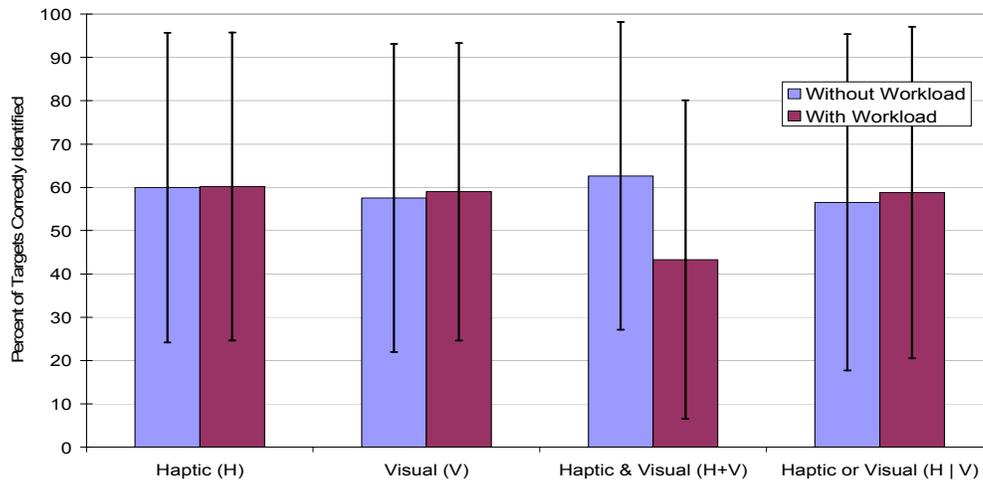
Figure 7: Visual target detection performance as a function of navigation signal type and presence of workload. In all figures, error bars represent standard deviations.
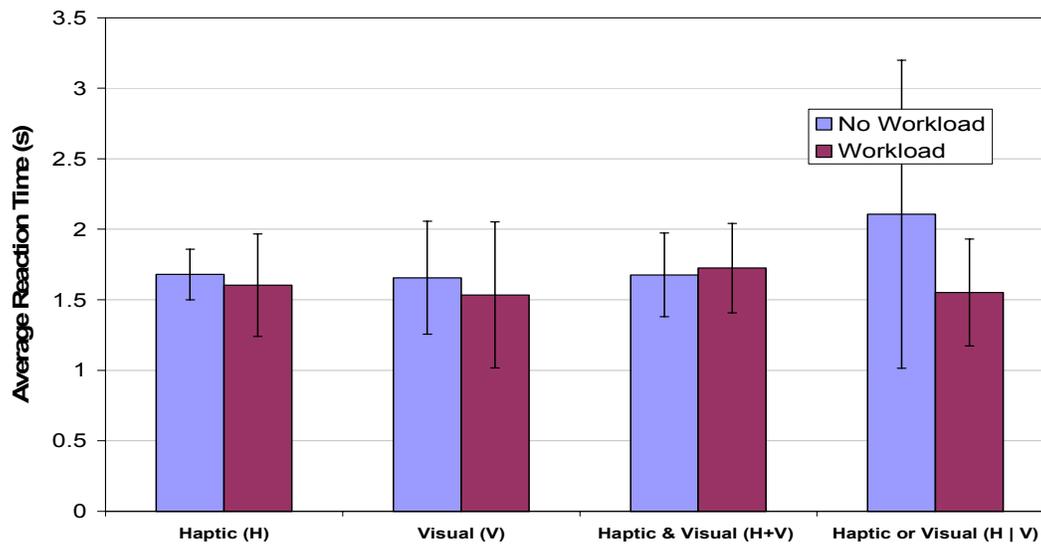
Figure 8 Visual target detection reaction times

types and workload levels in visual target acquisition time (average time between the presentation of the visual targets and theparticipant's pedal press response, as shown in Figure 8.

5.2 Navigation Task

We measured the rate of correct turns, using errors as an indication of the difficulty to notice and respond to navigation signals under the experiment conditions (Figure 9).

The signal-to-noise ratios for haptic and visual signals were previously calibrated to produce an 80%-correct workload-free navigation performance level. Actual performance in the corresponding experimental conditions (H_{WL-} and V_{WL-})

was indeed close to this: responses for H_{WL-} were correct, on average, for 24.4 / 30 (81.3%) trials (std dev = 5.3, range =13-30), and for V_{WL-} , 24.8 / 30 (82.6%; stdev = 2.8, range=20-30). Learning over the 8 experiment blocks also did not appear to be a factor, as performance in the H_{WL-} and V_{WL-} blocks did not depend on their sequence number. Thus, it appears that our calibration held throughout the experiment trials.

A two-factor repeated measures ANOVA (4 signal types x 2 levels of workload) confirmed significant main effects of workload (F=6.208, p = 0.028, partial η^2 = 0.341) and type (F = 4.016, p = 0.039, partial η^2 = 0.641). Post-hoc pairwise comparisons indicated significantly more correct turns in the reinforced condition (H+V) than following visual signals in the mixed condition ($H|V_{post-V}$) (p = 0.034). This trend held true without workload: there were significantly more correct turns for both the V_{WL-} (p = 0.018) and reinforced ($H+V_{WL-}$, p = 0.041) conditions than following visual signals in the mixed ($H|V_{post-V,WL-}$) condition. The corresponding comparisons with haptic signals in the $H+V_{post-H, WL-}$ condition were not significant.

It is useful to look more closely at the impact of workload on navigation response to unimodal signals. When only visual signals were used, results showed that the addition of the workload task affected people's ability to respond correctly (82.6% without workload vs. 74.1% with workload; $F_{1, 12} = 6.48$, p=0.026). In contrast, when only haptic signals were given, the addition of the workload task did not significantly affect people's ability to respond correctly to those haptic signals (81.3% without workload vs. 79.7% with workload). Unexpectedly, we observed a performance decrease with workload in the reinforced condition which was also significant in the multivariate analysis ($F_{1, 12} = 6.86$, p=0.022).

Thus, overall we found evidence that adding cognitive workload makes it more difficult to respond correctly to navigation signals when all of the signals are solely visual (V), as well as when all of the signals are multimodally reinforced (H+V). We did not find any evidence that cognitive workload affected navigation performance when the signals were solely haptic (H). In general, it is easier to respond correctly to reinforced signals than visual signals when presented alone or in the mixed condition (when they are unexpected).
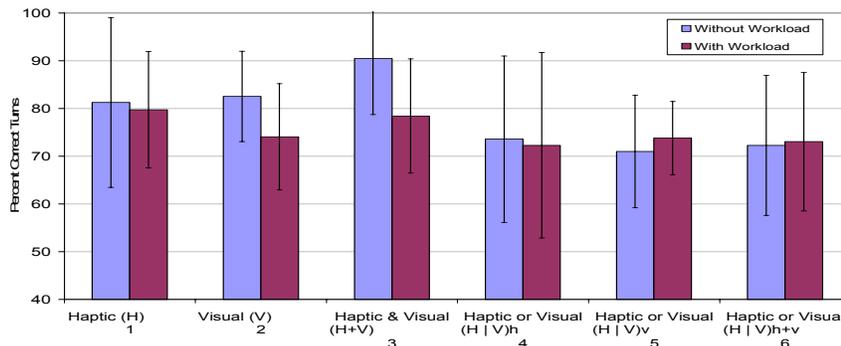
Figure 9: Navigation performance (percent correct turns) as a function of signal type and workload. Column pairs 4, 5 and 6 represent the results for the mixed condition. Column pair 4 represents the results for haptic signals in the mixed condition (H|V)h. Column pair 5 represents the results for visual signals in the mixed condition (H|V)v . Column pair 6 represents the results for both haptic and visual signals in the mixed condition (H|V)h+v.

5.3 Confidence

A repeated-measures ANOVA indicated significant main effects of cognitive workload (F = 10.232, p = 0.008, partial η^2 = 0.482) and type (F = 7.625, p = 0.008, partial η^2 = 0.718) on participant's estimates of how often they knew which way to turn. Post hoc pairwise comparisons indicated significant differences in confidence between the H and H|V conditions (p = 0.005), and between the H+V and H|V conditions (p = 0.007). Consistently with performance results, people reported that they knew which way to turn more often with haptic (H) or reinforcing (H+V) signals than with mixed signals (H|V).

Without workload, participants reported knowing which way to turn significantly more often in the H and H+V conditions than in the H|V condition (p = 0.006 and p = 0.004 respectively). Again, these findings reflect differences in performance.

Results indicated a significant effect of workload on confidence in the reinforced condition ($F_{1, 11}$ = 8.848, p=0.013). Thus in the reinforced condition, adding workload seemed to decrease reported knowledge of which way to turn.

Although not a statistically significant difference, confidence was higher in the haptic no-workload condition than in the visual no-workload condition, even though correct turn performance was the same (81.3% and 82.6%, respectively) in these two conditions (by calibration). Participants reported that they knew which way to turn 65% of the time for haptic no-workload (which suggests 81.5% performance when a 50% chance of guessing correctly for each turn is taken into consideration) and 53% of the time for visual no-workload (which suggests 73.5% performance).
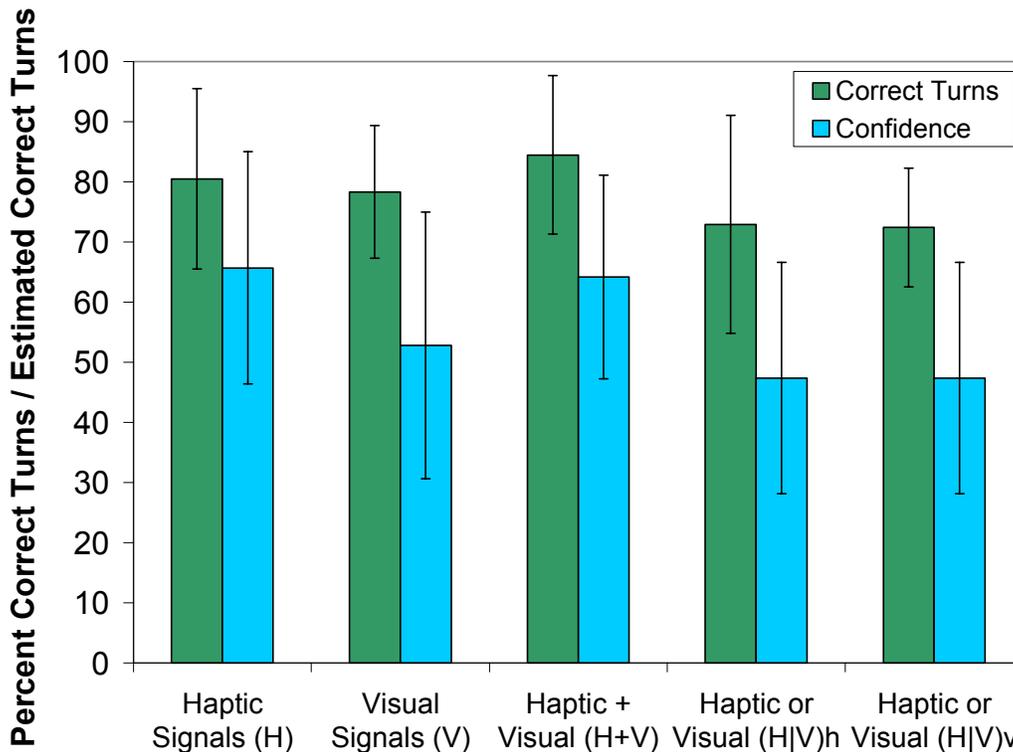
Figure 10 Performance and confidence across conditions

5.4 Participant Opinions

A majority of participants found the haptic signals more comfortable; a majority also thought it was more difficult to pay attention to visual signals compared to haptic signals while counting sentences. A detailed breakdown of questionnaire responses can be seen in Figure 11.

6 Discussion

The goal of this experiment was to make a comparison between the use of visual versus haptic signals to communicate information to a user engaged in a primary visual task, and to investigate the impact of cognitive workload on performance in these cases. This is an important issue because it will allow the design of user interfaces that communicate reliably without causing unnecessary distraction in the common context of primary visual tasks.

We started out asking the following questions:
1. Are visual and haptic navigation signals equally susceptible to the negative effects of cognitive workload during a visually-based search and response task?
2. Are visual and haptic navigation signals equally intrusive to a visual search and response task? (Do visual and haptic navigation signals cause different levels of interruption of the visual stimulus detection task?)?

We investigated these questions using a primarily visual spatial navigation paradigm, because it shares features with navigation during driving or piloting a plane. Our visual target detection task was intended to simulate some aspects of the visual attention needed to drive or fly.

6.1 Impact of Workload on Navigation Performance with Unimodal Cues

We hypothesized that in this context, visual navigation signals would be more adversely affected by the addition of a workload task than haptic navigation signals, because the visual system was already occupied by the visual target detection task. Our results support this idea: under workload, participants made more navigation errors when the navigation signals were visual compared to when the navigation signals were haptic.

That is, haptic navigation signals on their own were robust in the face of workload with no significant change in navigation performance; whereas visual signals alone were not robust, exhibiting a significant, 8.5% performance reduction.

The results obtained by Engstrom et al. [3], showed that hit rates for visual signal detection dropped (significantly) from 90% to 75% with the introduction of a (phone) workload task. The same study showed that hit rates for tactile signals dropped from 95% to 82% for the same conditions. It should be noticed that the hit rates for both tactile and visual signals were not equalized at any baseline for their experiment.

6.2 Impact of Workload on Performance with Reinforcing Multimodal Cues

Without workload, reinforced multimodal signals (H+V) improved navigation performance over the average for unimodal conditions H and V (82% to 91%, a difference of 9%); but an interesting effect occurred when workload was introduced to the H+V condition. Reinforcing multimodal navigation signals were *not* robust. The addition of cognitive workload caused a mean navigation performance decrease of 12% in the reinforced condition, but only a mean decrease of 4% (2% for haptic and 8% for visual) in the unimodal conditions. Stated another way, the benefit due to reinforcement disappeared when workload was added, with performance returning to unimodal levels.
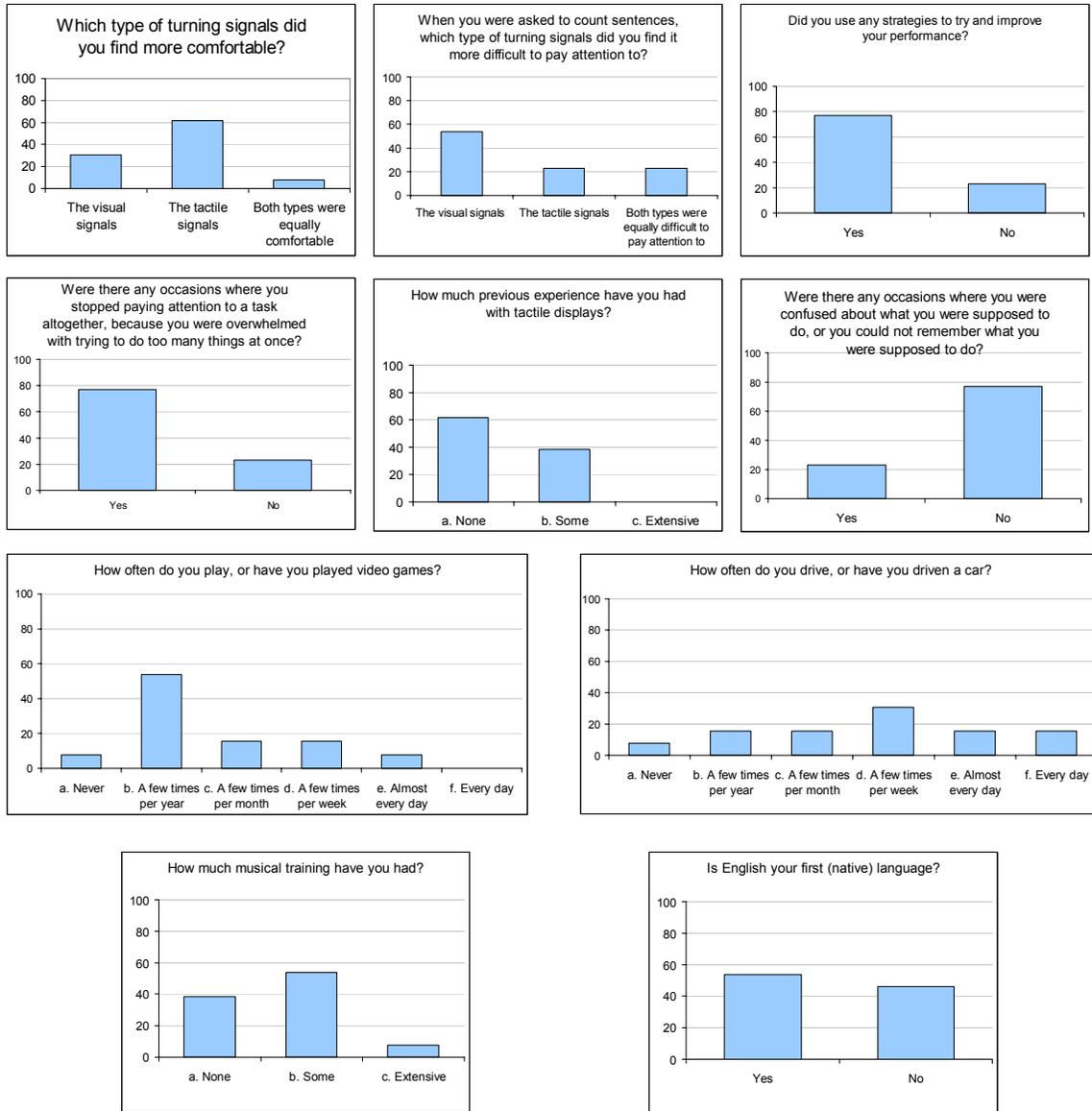
Figure 11 Breakdown of questionnaire responses

We suggest that this result may be due to an "overload" effect, where more cognitive resources are needed to process information from both modalities at the same time, due to the "loudness" of the incoming information. This effect may only be apparent when cognitive workload is added because the workload increases the total demand for cognitive resources past a threshold, so that all tasks can no longer be carried out easily at the same time.

6.3 Impact of the Navigation Task on Visual Search Performance

With respect to the question of intrusiveness, we predicted that visual navigation signals might be more attentionally intrusive than haptic cues, because the primary task being interrupted is visual. However, we found that the signal modality itself did not have an effect on participants' ability to detect the visual targets or their reaction

time in detecting the targets. Participants were told that the visual target detection task was the most important, so the lack of any difference in measured intrusiveness of haptic and visual navigation signals may simply reflect good ability to preserve performance in that task at the expense of performance in the navigation task.

We unexpectedly found that visual target detection rates dropped significantly when workload was added in the reinforced condition. We suggest that this may be another manifestation of the distraction effect described above. Presenting both signals plus workload seemed to increase cognitive demands beyond a threshold, causing performance to suffer in one or more tasks. At this point some participants may have been unable to maintain performance in the visual target detection task – cognitive demands were high enough that this performance suffered along with performance in the navigation task.

6.4 Important Design Implications

The findings presented in this paper have several important implications for the design of effective and safe user interfaces. In a visually demanding situation, haptic signals may be well-suited to deliver navigation information reliably, even when cognitive workload is high. Furthermore, participant feedback suggests that in this context, haptic signals are comfortable and intuitive (subjectively easier to attend to) compared to visual signals.

It is also important to be aware of problems that may arise due to cross-modal integration when two signals are given at once. Presenting both visual signals and haptic signals at the same time increases cognitive demand more than presenting either signal alone; when dual-modality signals are in use, the addition of non-visual workload raises cognitive demands and impairs detection of visual targets.

7 Conclusions and future work

In this experiment, detection of visual targets that appeared at random locations on the screen was the primary task. Haptic and visual navigation signals directed participants through a maze; participants found the two types of navigation signals equal in utility without additional cognitive workload. Adding cognitive workload made it harder to respond correctly to the visual navigation signals while the haptic signals proved immune, presumably because visual attention was already occupied by the visual target detection task. This is supported by the idea that there are separate pools of attention for each modality [7].

Multimodal reinforcement followed typical patterns of improved signal response in the absence of workload. However, with workload an important threshold seems to have been crossed; the multimodal signal appeared to have distracted the participants, making it more difficult to maintain performance of both the navigation task and the visual target detection task.

Together, these findings suggest that haptic signals can be a *more* robust, intuitive and subjectively preferred way to communicate navigation information to a user in a predominantly visual task than are visual signals, without being any more intrusive than a visual signal. Further, we posit that reinforcing multimodal cues should be used with caution in attentionally demanding contexts given their possibly deleterious effects.

Potential applications for haptic cues

of this type include most real-time, safety critical environments with a continuous guidance interface component, including automotive and aircraft systems [18]. A growing niche in the consumer world is for pedestrian navigation systems in handheld mobile devices, such as that described in [19]. It is important to note that while we employed one type of directional haptic cue (spatial distributed tactile stimuli), similar results might apply to a wide variety of other haptic stimuli and devices.

Several issues which will further inform the design of optimal haptic guidance systems need to be further investigated. In one direction, the investigations of the current study would fruitfully be extended to increasingly realistic instantiations of the target environment, and eventually to actual driving with a real-time navigational aid. More generally, we believe that haptic communication can be employed for more complex information transfers than the binary directional cues examined here. We do not yet know how much information can be encoded in a single haptic message, although a number of studies have explored the design and application of haptic icons [20-22]. In part, this is due to users' relative unfamiliarity with rich haptic signals. The new evidence of haptic signals' relative robustness to workload presented here is encouraging, but how will more complex signals fare? These questions and others are driving ongoing research by our group.

8 References

1. Rupert, A.H., *An instrumentation solution for reducing spatial disorientation mishaps: a more 'natural' approach to maintaining spatial orientation.* IEEE Engineering in Medicine and Biology, 2000: p. 71-80.
2. Van Veen, H.A.H.C.v.E., J.B.F. *Providing directional information with tactile torso displays.* in *Eurohaptics 2003.* 2003.
3. Engstrom, J., Aberg, N., & Johansson, E. *Comparison between visual and tactile signal detection tasks applied to the safety assessment of in-vehicle safety information systems.* in *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.* 2005.
4. Patten, C., Kircher, A., Ostlund, J., & Nilsson, L., *Using mobile telephones: cognitive workload and attention resource allocation.* Accident Analysis & Prevention, 2004. **36**(3): p. 341-350.
5. Wickens, C.D., *Engineering Psychology and Human Performance.* 1984.
6. Wickens, C.D., Hollands, J.G., *Engineering Psychology and Human Performance, 3rd ed.*, ed. P. Hall. 2000, Saddle River, NJ.
7. Alais, D., Morrone, C. & Burr, D., *Separate attentional resources for vision and audition.* J. Neuroscience, 2004. **1-12**.
8. Tan, R.G.a.H.Z., *Dynamic and Predictive Links between Touch and Vision.* Experimental Brain Research, 2002. **145**: p. 50-55.
9. Macaluso, E., & Driver, J., *Spatial attention and crossmodal interactions between vision and touch.* Neuropsychologia, 2001. **39**: p. 1304-1316.
10. Stein, B.E., M.A. Meredith, *THE MERGING OF THE SENSES.* 1993, Cambridge, MA: The MIT Press.
11. MIURA, T., *Active function of eye movement and useful field of view in a realistic setting,* in *From Eye to Mind: Information Acquisition in Perception, Search, and Reading,* d.Y.a.R.P. R. Groner, Editor. 1990: Amsterdam: North-Holland. p. 119 -127.
12. Rantanen, E.G., J., *The effect of mental workload on the visual field size and shape.* Ergonomics, 1999. **42**(6): p. 816-834.
13. H. Tan, A.L., and R. Traylor. *A Psychophysical Study of Sensory Saltation with an Open Response Paradigm.* in *9th Int'l Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems.* 2000: Am. Soc. Of Mechanical Engineers, Dynamic Systems and Control Division.
14. Baldwin, C.L.C., J.T. *Mental workload as a function of traffic density: comparison of physiological, behavioural, and subjective*

indices. in *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.* 2003. Iowa City: University of Iowa.

15. www.tactaid.com.

16. Potter, *Short-term conceptual memory for pictures.* Journal of experimental psychology: Human learning and memory, 1976. **2**(5): p. 509 -522.

17. Taylor, M.M., Creelman, C. D., *PEST: Efficient Estimates on Probability Functions.* Journal of the Acoustical Society of America, 1967. **41**(4A): p. 782-787.

18. Enriquez, M.J., MacLean, K. E. *Impact of Haptic Warning Signal Reliability in a Time-and-Safety-Critical Task.* in *12th Annual Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems, IEEE-VR2004.* 2004. Chicago, USA.

19. Luk, J., Pasquero, J., Little, S., MacLean, K., Hayward, V., Levesque, V. *Haptics as a Solution for Mobile Interaction Challenges: Initial Design Using a Handheld Tactile Display Prototype.* in *ACM Conference on Human Factors in Computing Systems, CHI '06.* 2006. Montreal, Canada.

20. Enriquez, M.J., MacLean, K. E., Chita, C. *Haptic Phonemes: Basic Building Blocks of Haptic Communication.* in *Eighth International Conference on Multimodal Interfaces (ICMI'06).* 2006. Banff, Alberta, Canada.

21. MacLean, K., Enriquez, M. *Perceptual Design of Haptic Icons.* in *Eurohaptics.* 2003. Dublin, Ireland.

22. Chan, A., MacLean, K., McGrenere, J. *Learning and Identifying Haptic Icons Under Workload.* in *World Haptics.* 2005. Pisa, Italy.