# 8 Results’ analysis and discussion

Throughout the experiment, three data sources were gathered from the participants, and this chapter will show their values, explain the process to analyze the data and discuss their results. This chapter is divided into two sections, each one related to one of the objectives:

* Evaluation of assistive device from a human factors’ perspective in a virtual environment;
* Comparison between BVI users and sighted users.

From this point, the data from the blind participants will be called the “Blind” sample and the data from the sighted participants will be called the “Sight” sample.

## 8.1 Evaluation of assistive device from a human factors’ perspective in a virtual environment

In this section it will be presented the discussion between the gathered data and the first goal of this experiment, “is it possible to evaluate and compare concepts of assistive device from a human factors’ perspective in a virtual environment? What are the main limitations of the use of a virtual reality environment?”. This discussion will be divided in two different subsections, one for the data from the questionnaires and the other for the data from the physiological sensors.

For the following ANOVA tests, a QQ plot and a residual plot was drawn to verify the residual normality. These Figures are in the Appendix F.

### 8.1.1 Data from questionnaires

There were 3 different questionnaires in this experiment. Each of these questionnaires was meant to verify one of the experiment goals:

* NASA-TLX;

Meant to verify the mental workload of the user. Is expected that after each “First” round, the mental workload would decrease and that one of the methods would have the least mental workload.

* Adapted SAGAT;

Meant to verify the situation awareness and the mental map of the user. Is expected to notice an increase from the “First” round to the “Return” round at each method.

* Guidance method’s questionnaire.

Meant to assess the user experience with each method.

8.1.1.1 NASA-TLX

It is possible to analyze the mental workload using NASA-TLX in two different ways. The first is by analyzing only the mental demand scale and the second is by analyzing the NASA-TLX score, which is an average of the scales’ rating.

* Analysis of the mental demand scale

The Table 8.7 presents the mental demand averages by each blinded participant on each scene and they are plotted in the Figures 8.5. The Figures 8.5 shows a systematic reduction on the perceived mental demand in all methods between the rounds. This shows that the participants started to get used with the device after the first use. The Figure 8.9 presents a box plot with the gathered mental demand. This Figure show that the results are rather similar. Also that the “Base” and “Audio” have the lesser mental demand and the presence of a haptic device elevated the mental demand.

TABLE 8.7 – Mental demand felled by the blinded participants.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Participant | Visual  Condition | Round |  |  |  |  |  |
| 001C | Blind | First | 3 | 1 | 14 | 3 | 6 |
|  |  | Return | 1 | 1 | 10 | 2 | 6 |
| 002C | Blind | First | 5 | 1 | 1 | 10 | 12 |
|  |  | Return | 1 | 1 | 1 | 10 | 3 |
| 003C | Blind | First | 5 | 5 | 5 | 8 | 1 |
|  |  | Return | 3 | 1 | 1 | 2 | 1 |
| 004C | Blind | First | 9 | 10 | 15 | 10 | 10 |
|  |  | Return | 7 | 10 | 14 | 8 | 10 |

The Table 8.8 shows the average mental demand in the “blind” sample and is possible to notice how the average perceived mental demand by the “blind” sample was lower during the “Audio” and the “Base” methods.

TABLE 8.8 – Mental demand average grouped by participant and visual Condition

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | 4.25 | 3.75 | 7.62 | 6.62 | 6.125 |

The Table 8.9 shows the mental demand variation between the rounds. This table shows that the mental demand variation from the “Virtual Cane” and the “Haptic Belt” were bigger than the other devices, but the “Base” method was still the biggest one. Maybe this happened because the users were more used to navigate using the traditional cane than with the other devices. This same data is plotted in the Figure 8.9.

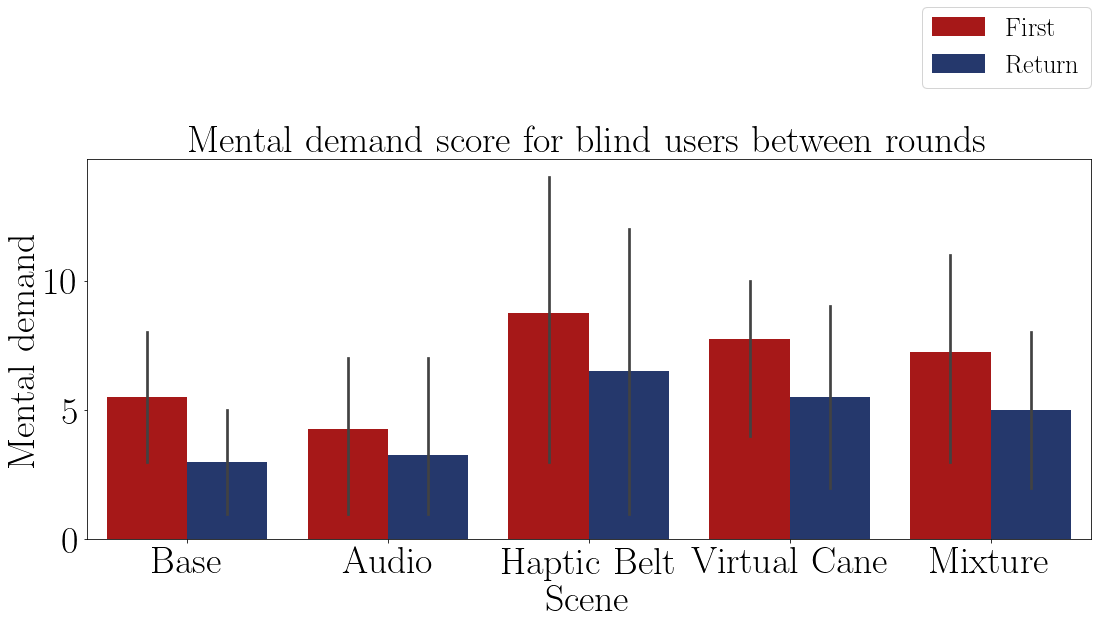


FIGURE 8.5 – Bar plot of the average mental demand of the blind participants on each method.

TABLE 8.9 – Mental demand variation grouped by participant and visual Condition

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | -52.2% | -20.0% | -28.8% | -32.1% | -18.8% |

The Tables 8.14 and 8.12 shows the ANOVA tests p-value of the mental demand of the “blind” sample between the guidance methods and both p-values indicate that there is at least one method that is statistically equal to one of the other methods.

The Table 8.15 presents the results of a pairwise Fisher LSD test of the blind mental demand average between all the guidance methods. The results show that only the “Audio” and the “Mixture” have a similar mental demand as the “Base” method.

The Table 8.13 presents the conclusion of a pairwise Fisher LSD test of the blind mental demand variation between all the guidance methods. The results show that only the “Virtual Cane” has a similar mental demand as the “Base” method.

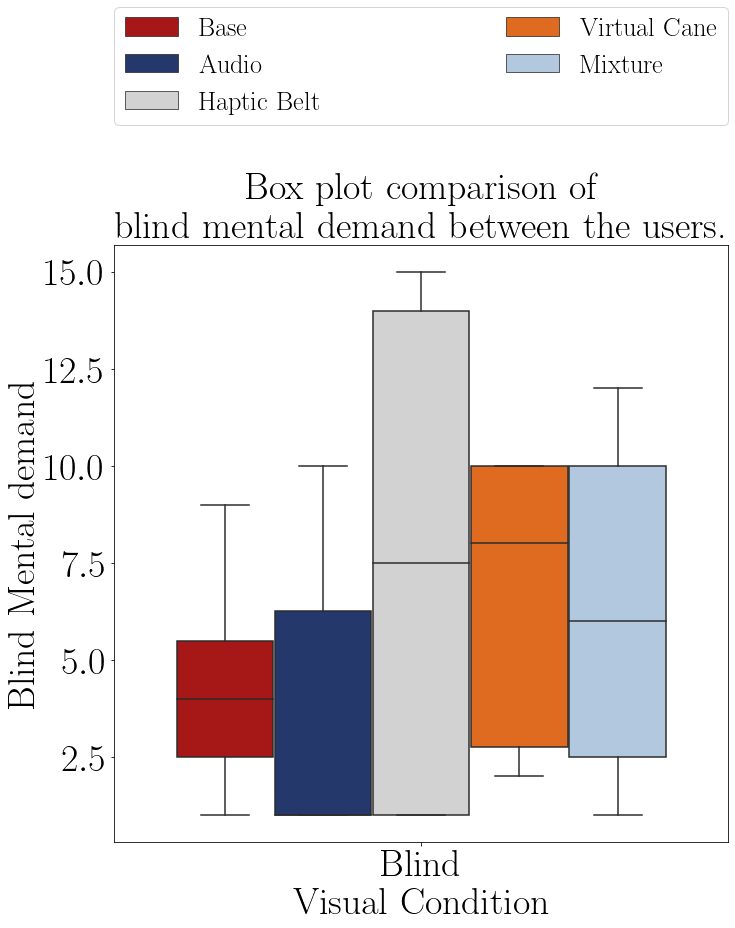


FIGURE 8.9 – Box plot of the mental demand from the blind participants of each method.

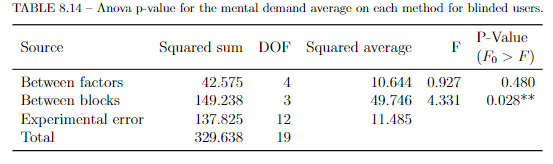


TABLE 8.12 – ANOVA p-value for the mental demand variation on each method for blinded users.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Squared sum | DOF | Squared average | F | P-Value  (*F*0 *> F*) |
| Between factors | 2901.806 | 4 | 725.451 | 0.683 | 0.617 |
| Between blocks | 5263.745 | 3 | 1754.582 | 1.652 | 0.230 |
| Experimental error | 12744.201 | 12 | 1062.017 |  |  |
| Total | 20909.752 | 19 |  |  |  |

To close up, according to the LSD test at Table 8.15 and the Tables 8.8, the average mental demand of the “Audio”, the “Mixture” are not statistically different from the “Base” so the average exposed on the Table 8.8 are not true, so there were no device that caused a lesser mental demand than the “Base” method. And, according to the LSD test at Table 8.13 and the Tables 8.9, the mental demand variation of the the “Mixture” is not statistically different from the “Base”, so the device that caused the biggest mental demand variation were the “Haptic Belt” and the “Virtual Cane”.

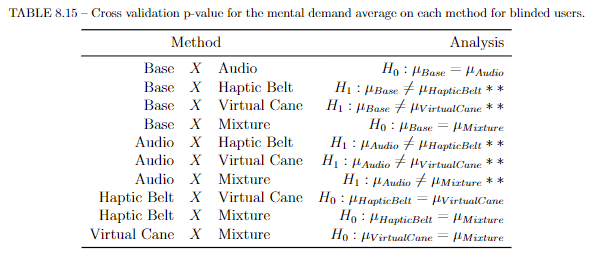


TABLE 8.13 – Cross validation p-value for the mental demand variation on each method for blinded users.

|  |  |  |  |
| --- | --- | --- | --- |
| Method | | | Analysis |
| Base | *X* | Audio | *H*1 : *µBase ≠* *µAudio* ∗∗ |
| Base | *X* | Haptic Belt | *H*1 : *µBase* *≠ µHapticBelt* ∗∗ |
| Base | *X* | Virtual Cane | *H*0 : *µBase* = *µV irtualCane* |
| Base | *X* | Mixture | *H*1 : *µBase ≠* *µMixture* ∗∗ |
| Audio | *X* | Haptic Belt | *H*0 : *µAudio* = *µHapticBelt* |
| Audio | *X* | Virtual Cane | *H*0 : *µAudio* = *µV irtualCane* |
| Audio | *X* | Mixture | *H*0 : *µAudio* = *µMixture* |
| Haptic Belt | *X* | Virtual Cane | *H*0 : *µHapticBelt* = *µV irtualCane* |
| Haptic Belt | *X* | Mixture | *H*0 : *µHapticBelt* = *µMixture* |
| Virtual Cane | *X* | Mixture | *H*0 : *µV irtualCane* = *µMixture* |

* Analysis of the NASA-TLX score

The Table 8.14 presents the NASA score averages by each blind participant on each scene and they are plotted in the Figures 8.10. It is noticeable that after each “First” round the NASA score diminishes. The Figure 8.14 has similar conclusion as Figure 8.9 for the mental demand. The scores are rather similar, the “Base” and the “Audio” method has smaller NASA score and the presence of a haptic device has risen the NASA score.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Participant | Visual  Condition | Round |  |  |  |  |  |
| 001C | Blind | First | 4.83 | 4.00 | 8.83 | 5.17 | 6.333 |
|  |  | Return | 4.17 | 4.00 | 6.67 | 4.50 | 6.167 |
| 002C | Blind | First | 6.33 | 4.83 | 4.83 | 9.00 | 7.000 |
|  |  | Return | 4.50 | 4.83 | 4.83 | 7.00 | 5.167 |
| 003C | Blind | First | 4.00 | 4.00 | 5.33 | 6.67 | 3.500 |
|  |  | Return | 4.00 | 3.83 | 3.67 | 3.50 | 3.500 |
| 004C | Blind | First | 9.83 | 10.00 | 12.67 | 9.67 | 11.000 |
|  |  | Return | 8.67 | 9.17 | 11.67 | 9.33 | 10.833 |

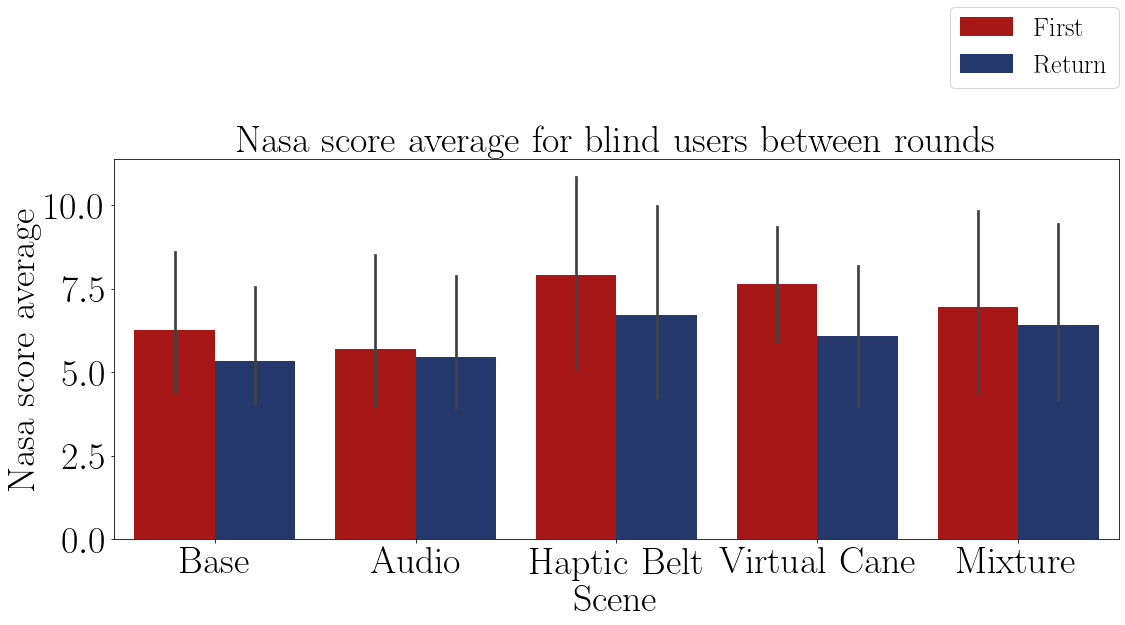


FIGURE 8.10 – Bar plot of the average Nasa-TLX score of the blind participants on each method.

The Table 8.15 shows the average NASA score grouped by visual condition and it also shows a similar pattern as the mental demand, so the lowest NASA score was in the “Audio” and “Base” methods.

TABLE 8.15 – NASA-TLX score grouped by participant and visual Condition

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | 5.79 | 5.58 | 7.31 | 6.85 | 6.688 |

The Table 8.16 shows the NASA score variation grouped by visual condition and it shows the difference between the rounds. The biggest variation was with “Virtual Cane” and with the “Haptic Belt”, similar as it was with the mental demand. The Figure 8.14 presents a barplot with the results from the Table 8.16.

TABLE 8.16 – NASA-TLX score grouped by participant and visual Condition.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | -13.7% | -3.1% | -15.9% | -21.5% | -7.6% |

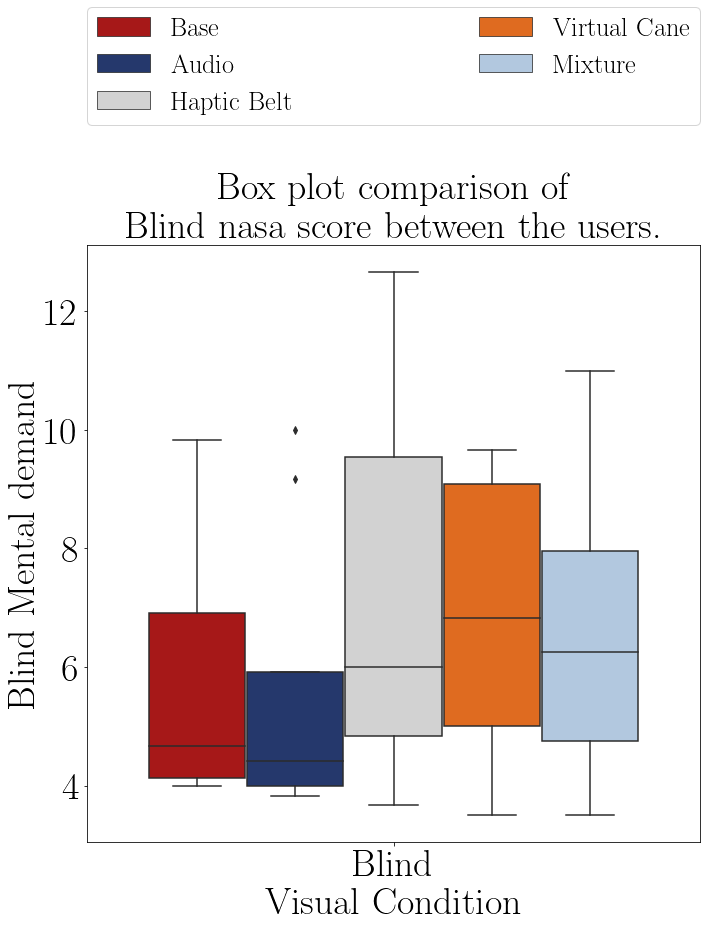
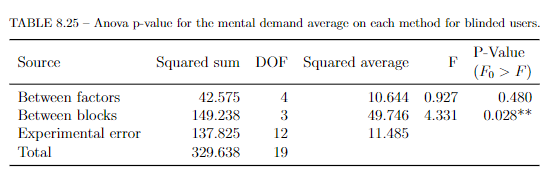
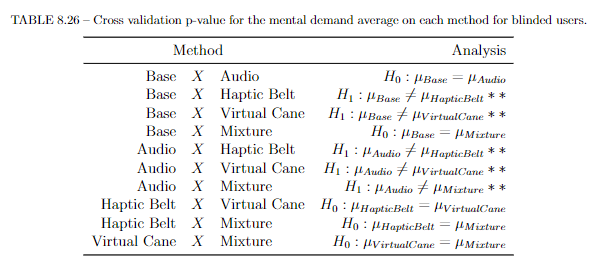


FIGURE 8.14 – Box plot of the NASA score of the blind participants of each method.

The Table 8.25 shows the ANOVA test p-value of the average NASA score of the “blind” sample between the guidance methods. The p-value indicates that there is at least one method that is statistically equal to one of the other methods so it’s recommended to do a pairwise analysis between all the methods.



The Table 8.26 presents the results of a pairwise Fisher LSD test of the blind NASA score average between all the guidance methods. The results show that the “Haptic Belt” and “Virtual Cane” proved different from the “Base” method, and both averages were bigger than the “Base” method.



The Table 8.19 shows the ANOVA test p-value of the NASA score variation between the rounds. The p-value indicates that there is at least one method that is statistically equal to one of the other methods so it’s recommended to do a pairwise analysis between all the methods.

TABLE 8.19 – ANOVA p-value for the NASA score variation on each method for blinded users.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Squared sum | DOF | Squared average | F | P-Value  (*F*0 *> F*) |
| Between factors | 825.191 | 4 | 206.298 | 1.065 | 0.415 |
| Between blocks | 315.471 | 3 | 105.157 | 0.543 | 0.662 |
| Experimental error | 2324.141 | 12 | 193.678 |  |  |
| Total | 3464.803 | 19 |  |  |  |

The Table 8.20 presents the results of a pairwise Fisher LSD test of the blind NASA score average between all the guidance methods. The results show that only “Audio” proved different from the ”Base” method.

TABLE 8.20 – Cross validation p-value for the NASA score variation on each method for blinded users.

|  |  |  |  |
| --- | --- | --- | --- |
| Method | | | Analysis |
| Base | *X* | Audio | *H*1 : *µBase ≠* *µAudio* ∗∗ |
| Base | *X* | Haptic Belt | *H*0 : *µBase* = *µHapticBelt* |
| Base | *X* | Virtual Cane | *H*0 : *µBase* = *µV irtualCane* |
| Base | *X* | Mixture | *H*0 : *µBase* = *µMixture* |
| Audio | *X* | Haptic Belt | *H*1 : *µAudio ≠* *µHapticBelt* ∗∗ |
| Audio | *X* | Virtual Cane | *H*1 : *µAudio ≠* *µV irtualCane* ∗∗ |
| Audio | *X* | Mixture | *H*0 : *µAudio* = *µMixture* |
| Haptic Belt | *X* | Virtual Cane | *H*0 : *µHapticBelt* = *µV irtualCane* |
| Haptic Belt | *X* | Mixture | *H*0 : *µHapticBelt* = *µMixture* |
| Virtual Cane | *X* | Mixture | *H*1 : *µV irtualCane ≠* *µMixture* ∗∗ |

According to LSD test at Table 8.25, the devices that caused a different score was the “Haptic Belt” and the “Virtual Cane”, but both methods resulted in a higher average, that means that a navigation with the presence of “Audio” guidance helps to easy the mental workload for the BVI users. And the LSD test at table 8.20 shows that the only variation that was different than the “Base” method was the “Audio” method, showing that BVI there were only a small learning in this method. That means that the users were already used in this method of navigation when not using their favorite, or most used, navigation method.

### 8.1.2 Adapted SAGAT

In this subsection, the SAGAT questionnaire is analyzed. Its result may give an idea of the mental map the participant is drawing. For each question a participant could score 1 point or a fraction of it. The total score of each blind participant is presented on the Table 8.25 and they are plotted in the Figures 8.15, where it is visually noticeable that the performance better the second time they visit the room. The Figure 8.16 shows that there are two groups of scores one with the “Base”, “Haptic Belt” and the “Mixture” methods, and the second group with the “Audio” and the “Virtual Cane” methods. The first group scored higher than the second one.

TABLE 8.25 – Adapted Sagat global score by participant and guidance method.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Participant | Visual  Condition | Round |  |  |  |  |  |
| 001C | Blind | First | 6.25 | 5.50 | 5.33 | 5.83 | 3.500 |
|  |  | Return | 6.25 | 6.50 | 8.50 | 5.50 | 5.500 |
| 002C | Blind | First | 6.75 | 4.50 | 3.99 | 4.50 | 6.250 |
|  |  | Return | 5.25 | 5.00 | 4.00 | 6.50 | 8.500 |
| 003C | Blind | First | 7.25 | 7.50 | 7.49 | 4.66 | 9.000 |
|  |  | Return | 10.00 | 10.00 | 8.50 | 9.00 | 9.000 |
| 004C | Blind | First | 7.50 | 6.00 | 7.66 | 4.99 | 6.500 |
|  |  | Return | 9.00 | 6.00 | 9.25 | 7.25 | 9.000 |

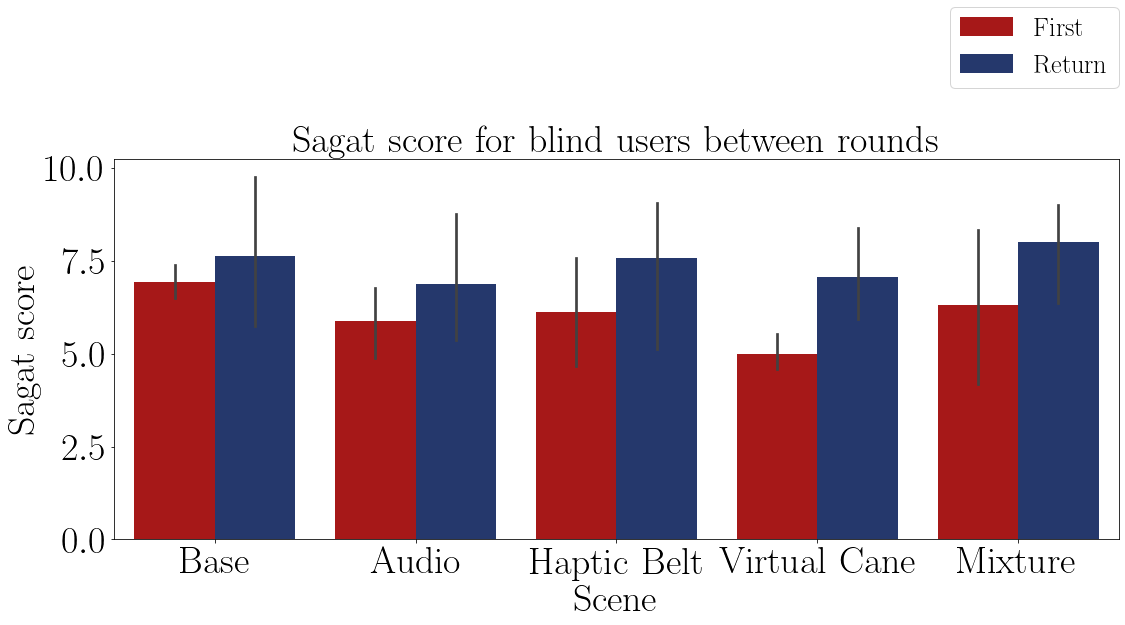


FIGURE 8.15 – Bar plot of the average SAGAT score of the blind participants on each method.

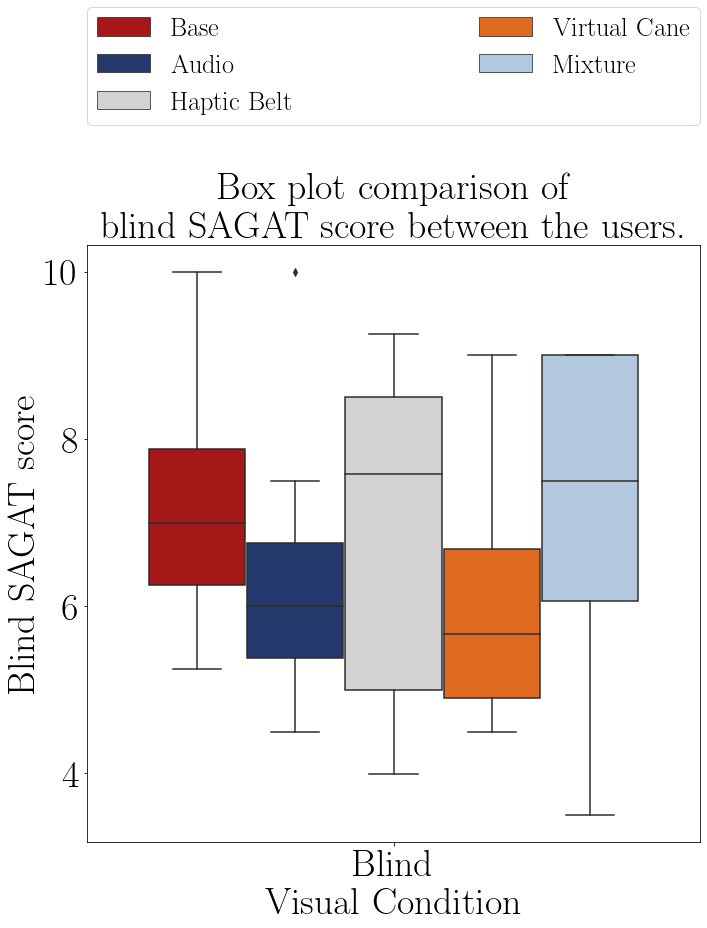


FIGURE 8.16 – Box plot of the SAGAT score of the blind participants on each method.

The Tables 8.26 and 8.27 shows the average and the variation between the rounds of the SAGAT score grouped by visual condition. The first table show that no method proved batter than the “Base” method and that the mixture of the devices was better than each one alone. The second table shows that the presence of a haptic method provoked a bigger variation on this score. This shows could be a consequence of the learning curve from the devices.

TABLE 8.26 – Adapted SAGAT score average grouped by participant and visual Condition.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | 7.28 | 6.38 | 6.84 | 6.03 | 7.156 |

TABLE 8.27 – Adapted SAGAT score variation grouped by participant and visual Condition

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | 8.9% | 15.7% | 23.5% | 44.3% | 32.9% |

The Table 8.30 shows the ANOVA test p-value of the Sagat score average of the “blind” sample between the guidance methods presented in the Table 8.25. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.

TABLE 8.30 – ANOVA p-value for the SAGAT score on each method for blinded users.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Squared sum | DOF | Squared average | F | P-Value  (*F*0 *> F*) |
| Between factors | 4.461 | 4 | 1.115 | 0.966 | 0.461 |
| Between blocks | 24.116 | 3 | 8.039 | 6.962 | 0.006\*\* |
| Experimental error | 13.856 | 12 | 1.155 |  |  |
| Total | 42.432 | 19 |  |  |  |

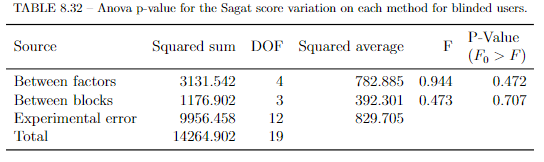
The Table 8.31 presents the analysis of a pairwise Fisher LSD test of the blind average SAGAT score between all the guidance methods. The results show that the “Audio” and ”Virtual Cane” caused a different SAGAT score than the one noticed on the ”Base” method.

The rest of the methods did not significantly change it.

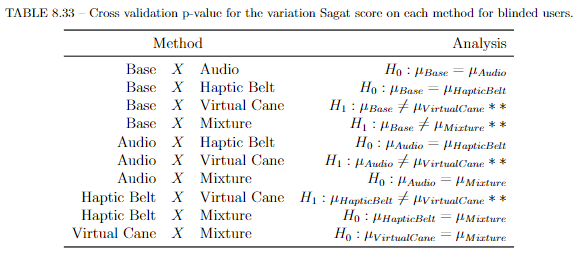
TABLE 8.31 – Cross validation p-value for the Sagat score on each method for blinded users.

|  |  |  |  |
| --- | --- | --- | --- |
| Method | | | Analysis |
| Base | *X* | Audio | *H*1 : *µBase ≠* *µAudio* ∗∗ |
| Base | *X* | Haptic Belt | *H*0 : *µBase* = *µHapticBelt* |
| Base | *X* | Virtual Cane | *H*1 : *µBase ≠* *µV irtualCane* ∗∗ |
| Base | *X* | Mixture | *H*0 : *µBase* = *µMixture* |
| Audio | *X* | Haptic Belt | *H*0 : *µAudio* = *µHapticBelt* |
| Audio | *X* | Virtual Cane | *H*0 : *µAudio* = *µV irtualCane* |
| Audio | *X* | Mixture | *H*1 : *µAudio ≠* *µMixture* ∗∗ |
| Haptic Belt | *X* | Virtual Cane | *H*1 : *µHapticBelt ≠* *µV irtualCane* ∗∗ |
| Haptic Belt | *X* | Mixture | *H*0 : *µHapticBelt* = *µMixture* |
| Virtual Cane | *X* | Mixture | *H*1 : *µV irtualCane ≠* *µMixture* ∗∗ |

The Table 8.32 shows the ANOVA test p-value of the SAGAT score variation of the “blind” sample between the guidance methods presented in the Table 8.25. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.



The Table 8.33 presents the analysis of a pairwise Fisher LSD test of the blind variation SAGAT score between all the guidance methods. The results show that The ”Virtual Cane” and the “Mixture” method caused a different SAGAT score than the one noticed on the ”Base” method. The rest of the methods did not significantly change it.



According to LSD test at Table 8.31 and the Figure 8.16, the devices that caused a different score was the “Audio” and the “Virtual Cane”, but both methods were the lowest SAGAT score. The LSD test at table 8.20 shows that the only variation that was different than the “Base” method was the “Audio” method, showing that BVI there were only a small learning in this method. That means that the users were already used in this method of navigation when not using their favorite, or most used, navigation method.

8.2.3 Guidance method’s questionnaire.

Finally, the Questionnaire is analyzed to give an idea about the impressions of the users with each device. This is an important evaluation to seek their impressions of each method. The higher the score, the more the user was satisfaction with that method. The Table 8.28 shows the average score of each method and they are plotted in the Figures 8.19 and the Figure 8.21 has a box plot of the answered values.

TABLE 8.28 – Guidance method questionnaire average score grouped by participant.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Participant | Audio | Haptic  Belt | Virtual  Cane | Mixture | Visual Condition |
| 001 | 0.75 | 0.49 | 0.57 | 0.69 | Sight |
| 001C | 0.77 | 0.54 | 0.63 | 0.87 | Blind |
| 002C | 0.86 | 0.74 | 0.54 | 0.93 | Blind |
| 003 | 0.76 | 0.54 | 0.54 | 0.78 | Sight |
| 003C | 0.93 | 0.57 | 0.54 | 0.74 | Blind |
| 004 | 0.86 | 0.60 | 0.79 | 0.76 | Sight |
| 004C | 0.88 | 0.49 | 0.40 | 0.73 | Blind |
| 005 | 0.61 | 0.57 | 0.75 | 0.84 | Sight |

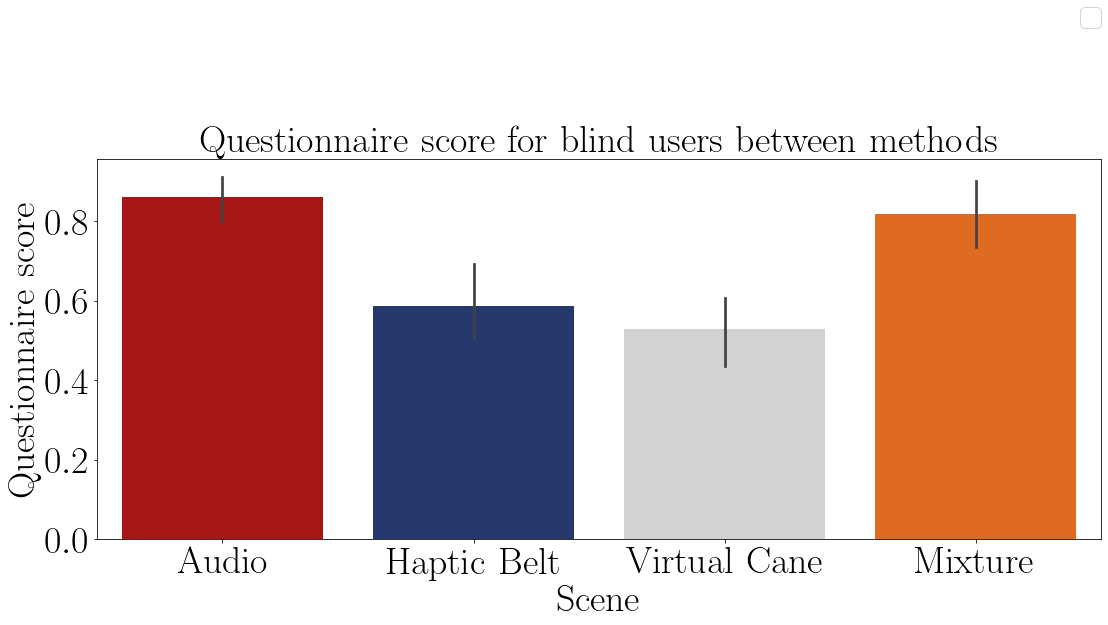


FIGURE 8.19 – Bar plot of the average mental demand of the blind participants on each method.

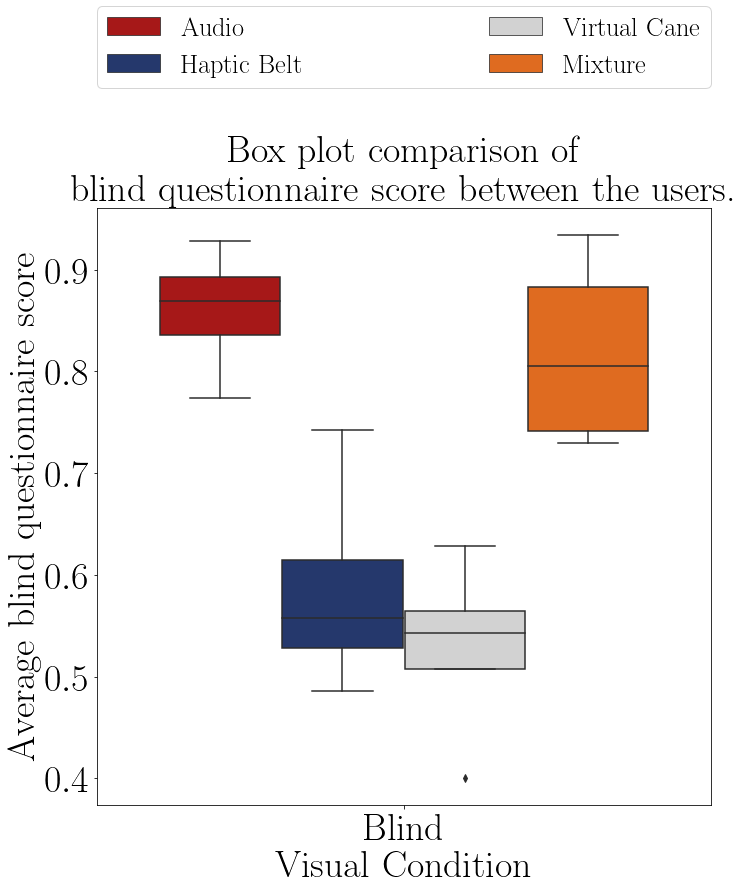


FIGURE 8.21 – Boxplot of the average questionnaire score of each group.

The Table 8.29 show the average questionnaire score of each participant and it appears that the presence of the “Audio” method brought more satisfaction to the user.

TABLE 8.29 – Guidance method questionnaire average score grouped by visual condition.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Visual Condition | Audio | Haptic Belt | Virtual Cane | Mixture |
| Blind | 0.86 | 0.59 | 0.53 | 0.82 |

The Table 8.31 shows the ANOVA test p-value of the Sagat score average of the ”blind” sample between the guidance methods presented in the Table 8.28. The p-value indicates that all scores are significantly different from each other. That means that the highest scores shown in Table 8.29, which are the “Audio” and the ”Mixture” methods were the most favorite by the blind participants, whilst the “Haptic Belt” and “Virtual Cane” were the most unfavorite devices. The participant did comment about those two last devices, saying that they were not precise enough, confusing and very different from what they are used to use.

TABLE 8.31 – ANOVA p-value for the questionnaire score on each method for blinded users.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Squared sum | DOF | Squared average | F | P-Value  (*F*0 *> F*) |
| Between factors | 0.329 | 3 | 0.110 | 15.677 | 0.001\*\* |
| Between blocks | 0.042 | 3 | 0.014 | 2.014 | 0.183 |
| Experimental error | 0.063 | 9 | 0.007 |  |  |
| Total | 0.434 | 15 |  |  |  |
|  |  |  |  |  |  |

### 8.1.2 Data from physiological sensors

There were 3 different sensors in this experiment, 2 that collected physiological data and the one left collected temperature. The last one was used to eliminate the temperature influence on the GSR sensor. These were all used to assess Mental Workload.

* Electrocardiogram (ECG) data;

Two features are extracted from the ECG, heartrate (BPM) and heartrate variance (SDNN).

Is expected that the heartrate increases at every “First” round and then a slight decrease in the next round. The heartrate variance is expected to decrease in the “First” round and a slight increase in the next round.

* Galvanic skin reaction and temperature data;

Is expected that the GSR average to increase at every “First” round and then a slight decrease in the next round.

### 8.1.2.1 Electrocardiogram (ECG) data

At the beginning of each experience, a baseline data was gathered to establish a comparison between the normal state of the user and the scenes’ induced state. After the data gathering, an algorithm in Python was used to read the data and separate it accordingly to each participant, method and round. The algorithm followed the steps above:

* Outliers remotion;

Since the participants moved during the whole experience a lot of noise was collected by the sensors

* Normalization between -1 and 1;
* Peak detection;

If the results were appropriate:

* + Heartbeat interval calculation
  + File save to be used in Kubius HRV Standard.

If the results were not appropriate:

* + Tune peak detection method’s parameters
  + Heartbeat interval calculation
  + File save to be used in the next software.

This judgment was made by analyzing the plotted ECG signal and the detected peaks. Kubios HRV Standard is a heart rate variability (HRV) analysis software for personal non-commercial use. The Kubios HRV Standard makes it possible to use your HR monitor to examine the health of the cardiovascular system or to evaluate stress and recovery (KUBIUS..., ). At Kubius, the file with the intervals was analyzed and the results were saved in a report file to be read in python again. Back in python the results were plotted, tabled and statistically tested as the other data. In Appendix D there is a diagram with a pseudo-algorithm of this process.

This analysis was made by comparing the baseline values with the values of each round individually and between the round values themselves.

8.3.1.1 Analysis of the heartbeat frequency (BPM)

The Table 8.32 presents the average heart rate by each blind participant on it’s baseline and on each scenes and they are plotted in the Figures 8.22. It is possible to see that the previous expectation was not seen. There was no heart rate increase by any participant with the exception only of the “sight” sample in the “First” round of the ”Base” method.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Baseline | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Part. | Visual  Condition | Round |  |  |  |  |  |  |
| 001C | Blind | First | 78.33 | 75.75 | 60.71 | 71.17 | 59.07 | 68.24 |
|  |  | Return |  | 71.05 | 58.61 | 66.22 | 64.20 | 70.76 |
| 002C | Blind | First | 67.78 | 48.69 | 38.67 | 48.74 | 46.89 | 52.23 |
|  |  | Return |  | 52.46 | 47.58 | 58.97 | 56.75 | 58.25 |
| 003C | Blind | First | 63.45 | 68.37 | 69.89 | 70.95 | 69.41 | 66.94 |
|  |  | Return |  | 67.34 | 67.44 | 69.68 | 68.82 | 67.37 |
| 004C | Blind | First | 78.30 | 75.09 | 73.55 | 73.70 | 71.94 | 74.03 |
|  |  | Return |  | 74.74 | 74.79 | 74.02 | 72.69 | 67.34 |

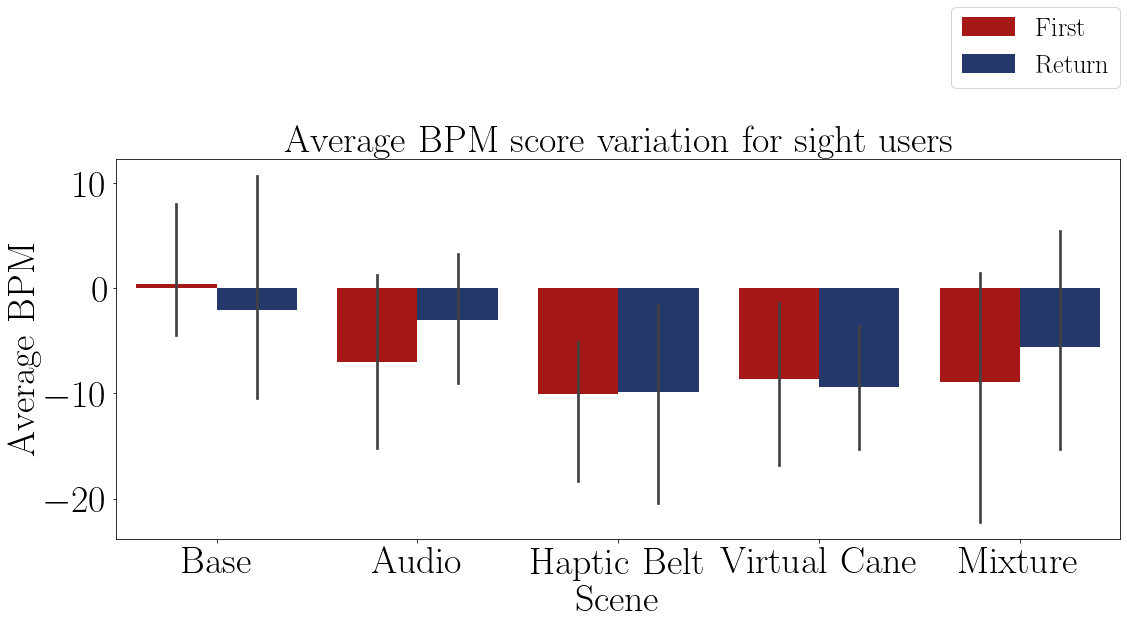


FIGURE 8.22 – Bar plot of the average heart rate of the blind participants on each method.

The Table 8.33 show the average heartbeat frequency variation between the rounds of each group. Despite all the variations being negative, it possible to see that the ”Audio” and the ”Virtual cane” provoked the highest variation in heartrate. That means that both devices caused the biggest mental workload relief between the rounds.

TABLE 8.33 – ECG average BPM average in relation to the baseline grouped by participant and visual Condition.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic Belt | Virtual Cane | Mixture |
| Blind | -7.3% | -14.3% | -6.9% | -10.9% | -8.4% |

The Figures 8.24 show a comparison between the methods and that it is possible to divide them in to two groups based on their similarity. One with “Base”, “Haptic Belt” and “Mixture” methods and the other with “Audio” and “Virtual Cane”. The first group apper to have a smaller variation than the second, but the second has lower values than the first, which could mean that the second group provoked a smaller mental workload.

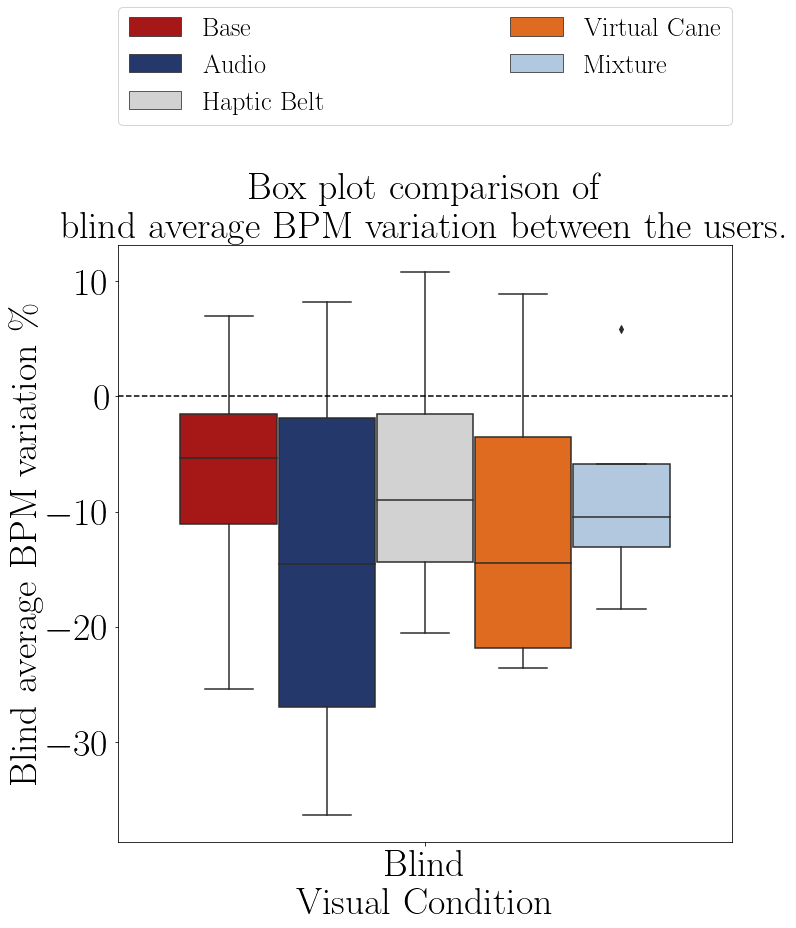


FIGURE 8.24 – Boxplot of the average heart rate of participants on each method.

The Table 8.36 shows the ANOVA test p-value of the heart rate frequency of the “blind” sample between the guidance methods presented in the Table 8.32. The p-value indicates that there is at least one method that is statistically equal to one of the other methods, and the LSD Fischer test in the Table 8.37 shows that the “Audio” and the “Virtual Cane” are different than the “Base” method.

TABLE 8.36 – ANOVA p-value for the BPM on each method for blinded users.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Squared sum | DOF | Squared average | F | P-Value  (*F*0 *> F*) |
| Between factors | 151.789 | 4 | 37.947 | 1.570 | 0.245 |
| Between blocks | 2934.674 | 3 | 978.225 | 40.471 | 0.000\*\* |
| Experimental error | 290.050 | 12 | 24.171 |  |  |
| Total | 3376.513 | 19 |  |  |  |

According to the ANOVA test at Table 8.36 and the LSD test at 8.37 the “Audio” and the “Virtual Cane” method provoked a different reaction than the “Base” method and analyzing the Table 8.33 and the Figure 8.24 both of them provoked the highest heartrate variation. That means that between the “First visit” and “Return” rounds they caused the biggest mental workload relief. But still, the negative variation between the baseline and the rounds was not expected as if the participants were not mentally provoked enough. This could be a consequence of the movement from the participants and procedures done to remove the out or that the tasks were not difficult enough in order to provoke a mental workload increase.

TABLE 8.37 – Cross validation p-value for the average BPM on each method for blinded users.

|  |  |  |  |
| --- | --- | --- | --- |
| Method | | | Analysis |
| Base | *X* | Audio | *H*1 : *µBase≠* *µAudio* ∗∗ |
| Base | *X* | Haptic Belt | *H*0 : *µBase* = *µHapticBelt* |
| Base | *X* | Virtual Cane | *H*1 : *µBase≠* *µV irtualCane* ∗∗ |
| Base | *X* | Mixture | *H*0 : *µBase* = *µMixture* |
| Audio | *X* | Haptic Belt | *H*1 : *µAudio≠* *µHapticBelt* ∗∗ |
| Audio | *X* | Virtual Cane | *H*1 : *µAudio≠* *µV irtualCane* ∗∗ |
| Audio | *X* | Mixture | *H*1 : *µAudio≠* *µMixture* ∗∗ |
| Haptic Belt | *X* | Virtual Cane | *H*1 : *µHapticBelt≠* *µV irtualCane* ∗∗ |
| Haptic Belt | *X* | Mixture | *H*0 : *µHapticBelt* = *µMixture* |
| Virtual Cane | *X* | Mixture | *H*0 : *µV irtualCane* = *µMixture* |

8.3.1.2 Analysis of the heartbeat variance (SDNN)

The Table 8.38 presents the standard deviation of the interval between heartbeat (SDNN) by each participant on their baseline and on each scene. They are plotted in the Figures 8.26 and it is possible to see that all of the users had an increase in the heartbeat variance. That means a relief in their mental workload.

TABLE 8.38 – ECG Average SDNN felled by the participants.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Baseline | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Part. | Visual  Condition | Round |  |  |  |  |  |  |
| 001C | Blind | First | 78.55 | 81.29 | 107.06 | 124.74 | 163.97 | 129.05 |
|  |  | Return |  | 120.72 | 130.88 | 131.59 | 157.59 | 124.79 |
| 002C | Blind | First | 93.77 | 73.76 | 98.86 | 81.14 | 33.98 | 79.29 |
|  |  | Return |  | 108.94 | 49.63 | 42.81 | 114.06 | 107.55 |
| 003C | Blind | First | 26.14 | 36.87 | 38.32 | 35.10 | 42.39 | 43.69 |
|  |  | Return |  | 52.75 | 41.20 | 44.26 | 42.60 | 46.14 |
| 004C | Blind | First | 20.98 | 70.73 | 86.83 | 62.56 | 85.90 | 70.47 |
|  |  | Return |  | 71.95 | 74.89 | 70.02 | 66.09 | 104.04 |

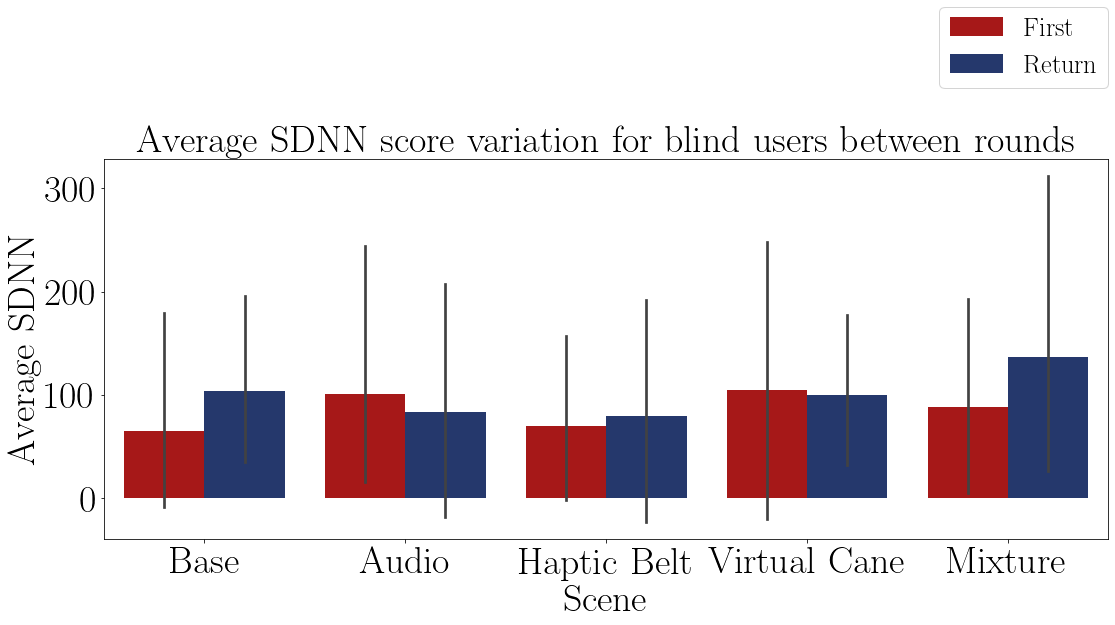


FIGURE 8.26 – Bar plot of the standard deviation of the heart of the blind participants on each method.

The Figures 8.28 shows all blind participants had a similar distribution of SDNN values.

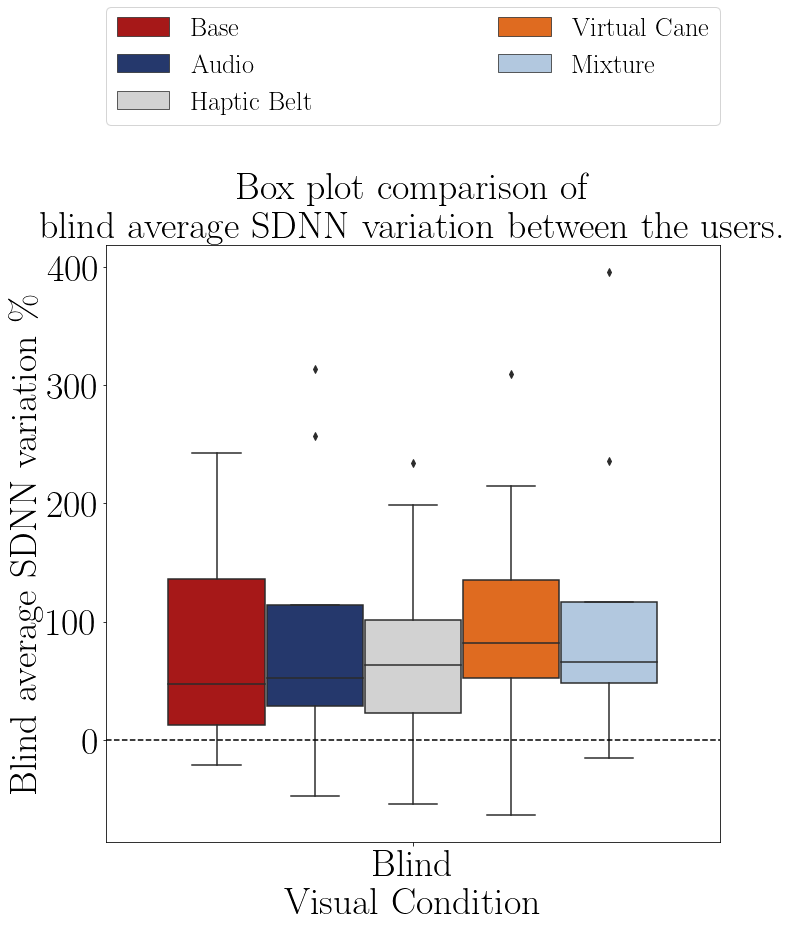


FIGURE 8.28 – Boxplot of the average heart rate of the participants on each method.

The Table 8.39 shows the variation of the heartbeat in each round of each group. In general, all the standard deviations increased, meaning that the mental workload decreased between the “Baseline” and the method.

TABLE 8.39 – ECG Average SDNN average in relation to the baseline grouped by participant and visual Condition.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic Belt | Virtual Cane | Mixture |
| Blind | 84.4% | 92.1% | 74.3% | 102.1% | 112.2% |

The Table 8.42 shows the ANOVA test p-value of the heart rate frequency of the “blind” sample between the guidance methods presented in the Table 8.38. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.

TABLE 8.42 – ANOVA p-value for the SDNN on each method for blinded users.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Squared sum | DOF | Squared average | F | P-Value  (*F*0 *> F*) |
| Between factors | 3519.680 | 4 | 879.920 | 1.564 | 0.247 |
| Between blocks | 214885.879 | 3 | 71628.626 | 127.314 | 0.000\*\* |
| Experimental error | 6751.365 | 12 | 562.614 |  |  |
| Total | 225156.923 | 19 |  |  |  |

The Table 8.43 presents the conclusion of a pairwise Fisher LSD test of the blind heart rate frequency variation between all the guidance methods. The results show that the ”Virtual cane” and the ”Mixture” method differs from the ”Base” method.

According to the ANOVA test at Table 8.42 and the LSD test at 8.43 and the Table 8.39 the “Virtual cane” and the ”Mixture” method did provoke an increase in the heartrate variance. That means that they caused the biggest mental workload relief in the participants. As was with the heartrate, this relief was not expected. It was expected that the SDNN decreased between the baseline and the method. The reasons that can explain this behavior are the same from the heartrate decrease: the movement from the participants or that the tasks were not mentally difficult enough

TABLE 8.43 – Cross validation p-value for the average SDNN on each method for blinded users.

|  |  |  |  |
| --- | --- | --- | --- |
| Method | | | Analysis |
| Base | *X* | Audio | *H*0 : *µBase* = *µAudio* |
| Base | *X* | Haptic Belt | *H*0 : *µBase* = *µHapticBelt* |
| Base | *X* | Virtual Cane | *H*1 : *µBase≠* *µV irtualCane* ∗∗ |
| Base | *X* | Mixture | *H*1 : *µBase≠* *µMixture* ∗∗ |
| Audio | *X* | Haptic Belt | *H*1 : *µAudio≠* *µHapticBelt* ∗∗ |
| Audio | *X* | Virtual Cane | *H*0 : *µAudio* = *µV irtualCane* |
| Audio | *X* | Mixture | *H*1 : *µAudio≠* *µMixture* ∗∗ |
| Haptic Belt | *X* | Virtual Cane | *H*1 : *µHapticBelt≠* *µV irtualCane* ∗∗ |
| Haptic Belt | *X* | Mixture | *H*1 : *µHapticBelt≠* *µMixture* ∗∗ |
| Virtual Cane | *X* | Mixture | *H*0 : *µV irtualCane* = *µMixture* |

### 8.1.2.2 Galvanic skin reaction.

The GSR analysis is made by analyzing its average and the accumulated value and comparing both features between the baseline and each round. The temperature was analyzed with the GSR to see if there is some influence and by a graphical analysis there was none. For the experiment, the GSR sensor was worn on the left hand for right-handed participant and on the right hand for left-handed participants.

8.3.2.1 Analysis of the GSR average

The Table 8.44 presents the average skin conductance by each participant on each scene and they are plotted in the Figures 8.30. It is possible to see that in all of the methods there was an increase in the average skin conductance, meaning that the user was aroused and maybe an increase in the mental workload. It also possible to notice that there were some repeated numbers.

TABLE 8.44 – Average GSR felled by the participants.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Baseline | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Part. | Visual  Condition | Round |  |  |  |  |  |  |
| 001C | Blind | First | 0.37 | 0.48 | 1.03 | 3.14 | 3.79 | 3.90 |
|  |  | Return |  | 0.83 | 1.58 | 2.81 | 4.04 | 4.57 |
| 002C | Blind | First | 0.17 | 0.91 | 0.23 | 0.17 | 0.17 | 0.17 |
|  |  | Return |  | 0.43 | 0.17 | 0.16 | 0.17 | 0.17 |
| 003C | Blind | First | 0.30 | 0.56 | 0.56 | 0.62 | 0.85 | 1.09 |
|  |  | Return |  | 0.62 | 0.63 | 0.65 | 0.92 | 1.06 |
| 004C | Blind | First | 1.24 | 2.34 | 3.07 | 3.49 | 2.28 | 2.23 |
|  |  | Return |  | 2.57 | 2.95 | 3.20 | 2.21 | 2.24 |

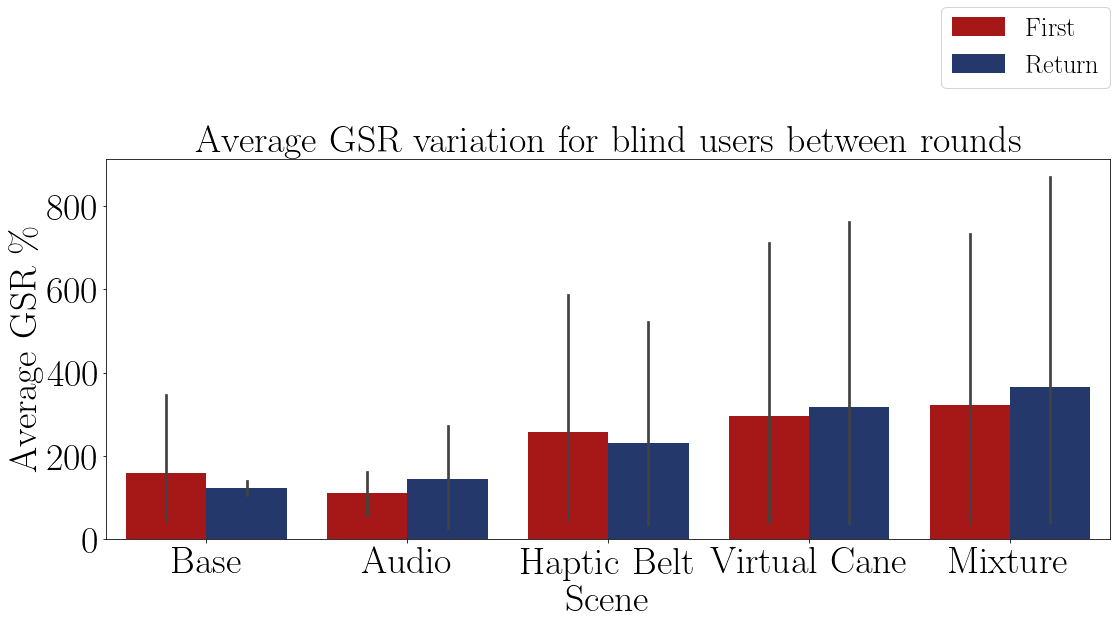


FIGURE 8.30 – Bar plot of the average skin conductance of the blind participants on each method.

The Figure 8.32 shows the distribution of the skin conductance variance and the Table 8.51 shows the average value of each distribution. It shows that the presence of a haptic device provoked an increase on the skin conductance. This means that with those devices the participants were more stressed, aroused or mentally overloaded during the use of these devices.

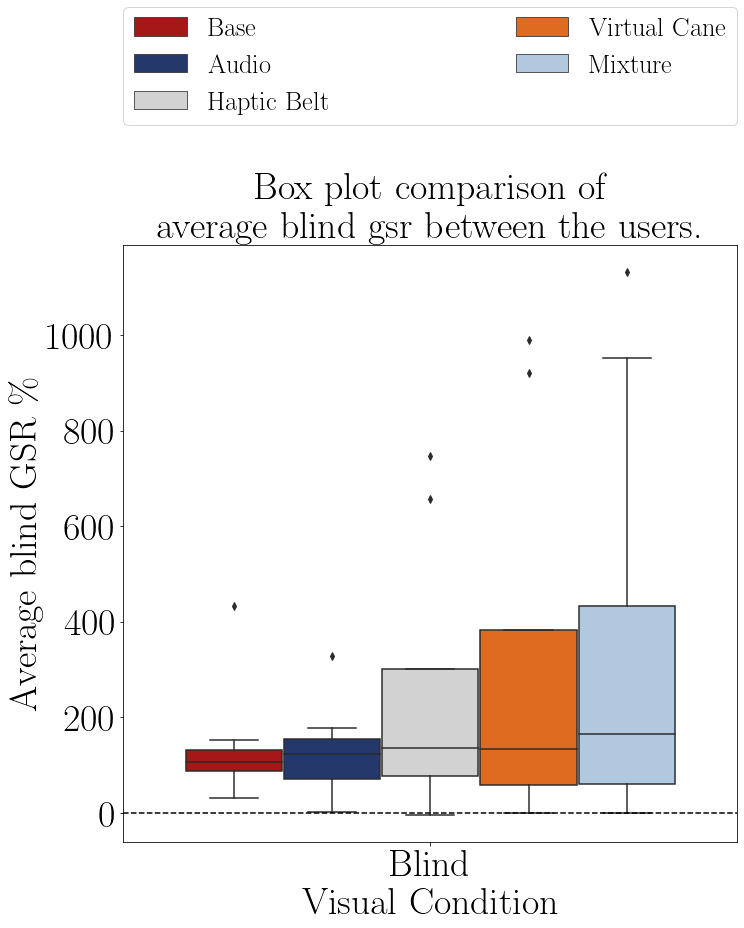
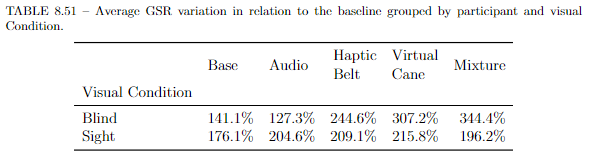


FIGURE 8.32 – Boxplot of the average skin conductance of the participants on each method.



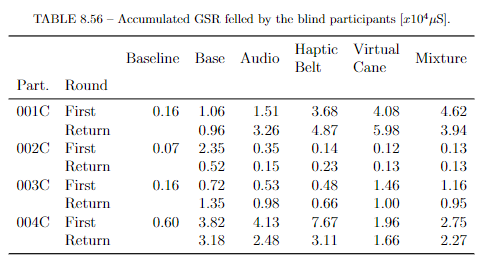
The Table 8.48 shows the ANOVA test p-value of the skin conductance variance of the “blind” sample between the guidance methods presented in the Table 8.44. The p-value indicates that all methods are different from each other. That means that the presence of a new device, especially haptic devices, provokes arousal or mental workload on the participants. This is comprehensible, since these devices were the ones that most of the participants complained about in the questionnaire.

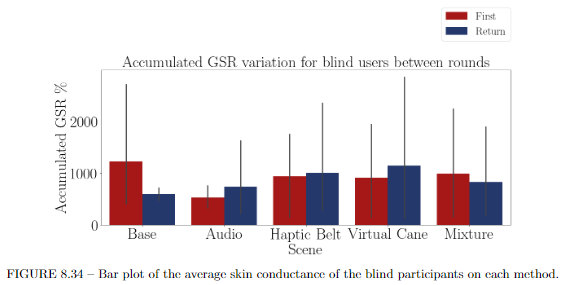
TABLE 8.48 – ANOVA p-value for the GSR score on each method for blinded users.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Squared sum | DOF | Squared average | F | P-Value  (*F*0 *> F*) |
| Between factors | 886616.269 | 4 | 221654.067 | 3.310 | 0.048\*\* |
| Between blocks | 1918983.649 | 3 | 639661.216 | 9.552 | 0.002\*\* |
| Experimental error | 803557.557 | 12 | 66963.130 |  |  |
| Total | 3609157.475 | 39 |  |  |  |

8.3.2.2 Analysis of the accumulated GSR

The Table 8.56 presents the accumulated skin conductance by each blind participant on each scene and their variation in relation to the baseline are plotted in the Figures 8.34. It is possible to see that in all of the methods there was an increase in the accumulated skin conductance, even between the “First” and “Return” rounds in the “Audio”, “Haptic Belt” and “Virtual Cane methods, meaning that the users were aroused and maybe an increase in the mental workload. There is also a big variation in the “Base” method.





The Figure 8.32 shows the distribution of the skin conductance distribution and the Table 8.51 shows the average value of each distribution. It shows that the presence of a haptic device provoked an increase on the accumulated skin conductance as well an increase of its variance. This means that with those devices the participants were more stressed, aroused or mentally overloaded during the use of these devices. The average of the “Base” method looks similar to the methods with the presence of haptic devices, but there is a high outlier inside that method, probably is average would be similar to the “Audio” method.

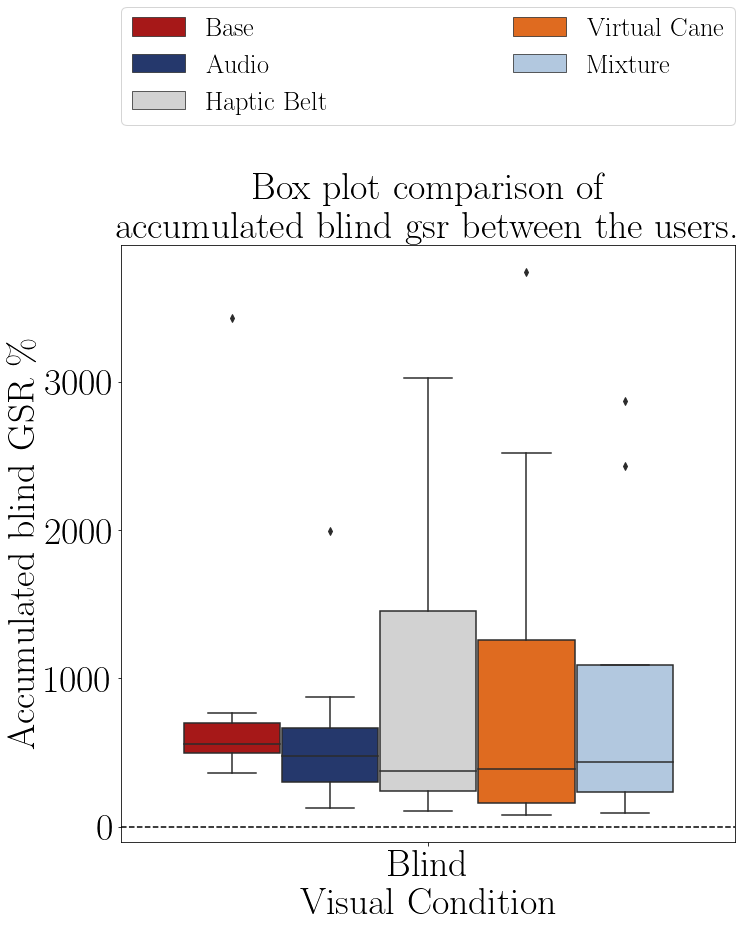
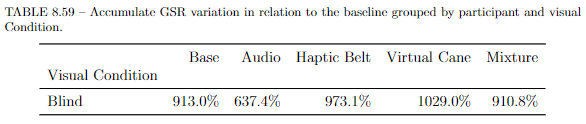
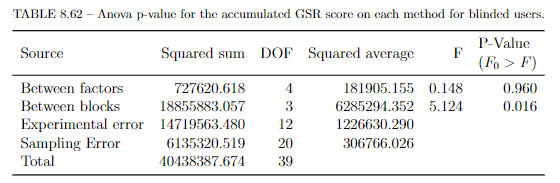
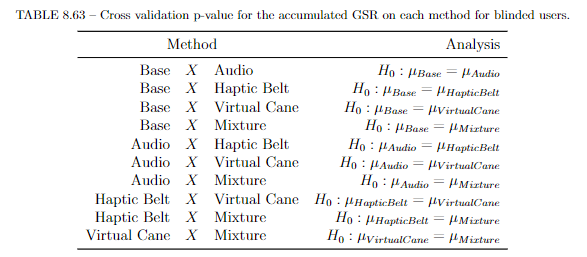


FIGURE 8.32 – Boxplot of the accumulated skin conductance of the blind participants on each method.



The Table 8.48 shows the ANOVA test p-value of the skin conductance variance of the “blind” sample between the guidance methods presented in the Table 8.44. The p-value indicates that probably there is no statistical difference between the methods and the LSD test on Table 8.63 confirms it. That means that all methods provoked the same level of stress or mental workload on the participants





## 2.4 Final remarks

The “Audio” method shown a higher performance among the other methods, and its presence increase the “Mixture” method performance as well. This probably happened because the participants are already used to use sound to guide themselves, especially environmental sounds. The environment sounds used inside the scenes that gave hints about locations where always the same (telephone ringing, laptop keyboard sounds, exterior noise, door opening and closing). It is likely that the participants felt more secure when it only had to focus on the sounds around him/her. This is reinforced by the fact that, during the “Audio” only guidance, half of the participants asked for none, or used only a few times the audio command option.

The fact that the haptic devices caused a higher average and a higher variation is probably due to the fact that the users had to learn and get used with them. Besides, for being just conceptual, their precision was not as big as they were expecting. That explains why their results were not as good as the “Base” or “Audio” methods and these results are correctly related to the satisfaction questionnaires, which scored them as the unsatisfied devices.

The ANOVA test and the boxplots most of the time had partially the same conclusion. In these cases the reason is the small sample size and the sample’s variety. The participants age range was from 26 to 56, with an average of 43.5 years, and the education range was from High School to 2 Graduations. That can impact in the user experience and as well in the questionnaires answer for the devices.

But all the participants showed a great enthusiasm before, during and after the research. They also recommend some modifications that would bring more realism for they. And of course, they made some complaints, such as:

* The speakers inside the HMD were not could enough for some to give them the precise location of its origin;
* The HMD was big enough to cover have of the participant’s face and that gave them a strange sensation, since some of them use the air or the wind feeling on the face to give them hints about the location of walls or other high obstacles;
* As said before, the precision of the vibration was not good for them to use the devices. That is mainly because of how the HMD position the user inside the virtual environment.

The user is represented as a vertical capsule, and the HMD is positioned on the top end of that capsule. If the user tilts his/her head down, as if they were facing the ground, the capsule rotates in relation to the HMD point making the virtual body of the user occupy a total different space from the reality. The Figure 8.33 represents that situation.

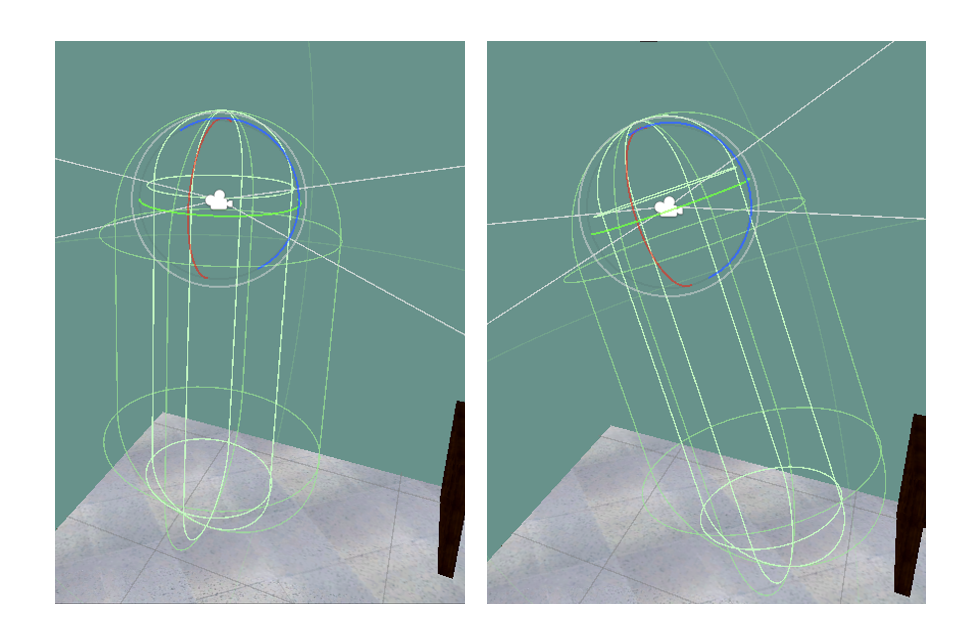


Figure 8.33 – Left: User with the head parallel to the ground. Right: User with the head slightly tilted to the ground.

* The vibration from the haptic belt was not intense enough sometimes.

8.2 Comparison between BVI users and sighted users.

In this section, the relationship between the second goal of this experiment, “do non-BVI users, when deprived from their vision, evaluate assistive devices in a similar way as BVI users?”, will be linked with the gathered data. As was the last section, this section will also be divided in the same subsections.

### 8.2.1 Data from questionnaires

Only the two questionnaires will be analyzed and it is expected that for:

* NASA-TLX;

There will be a noticeable difference between the sight sample mental workload and the blind sample mental workload.

* Adapted SAGAT;

Is expected to notice a difference between the “blind” sample and the “sight” sample.

8.1.1.1 NASA-TLX

* Analysis of the mental demand scale

The Table 8.8 presents the mental demand average on each scene grouped by visual condition. This tables shows a clear difference between the mental demand from the users in the different conditions. Something similar is also presented in the Table 8.9, where it shows the mental demand variation and that this variation is higher in the “blind” sample than in the “sighted” sample

TABLE 8.8 – Mental demand average grouped by participant and visual Condition

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | 4.25 | 3.75 | 7.62 | 6.62 | 6.125 |
| Sight | 4.00 | 11.75 | 14.25 | 9.38 | 12.125 |

TABLE 8.9 – Mental demand variation grouped by participant and visual Condition

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | -52.2% | -20.0% | -28.8% | -32.1% | -18.8% |
| Sight | -21.9% | -1.1% | -10.0% | -22.0% | -10.4% |

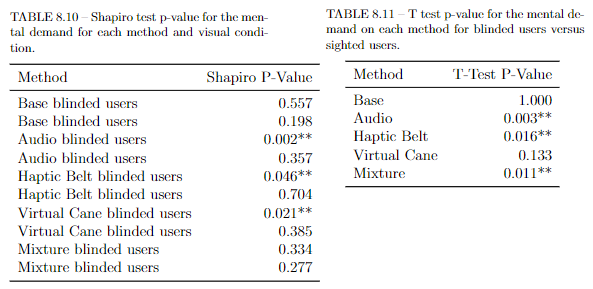
The Figure 8.7 presents a box plot of the mental demand scores of both groups plotted. A similar graphic is plotted in the Figure 8.8, where the average mental demand of both groups are plotted. In both figures it is noticeable that there is a difference between those two groups, but this difference is only statically meaningful if a hypothesis confirms it.

|  |  |
| --- | --- |
|  |  |
| FIGURE 8.7 – Boxplot of the average mental demand of participant. | FIGURE 8.8 – Barplot of the average mental demand of each group. |

The Shapiro–Wilk normality test on the Table 8.10 shows only the “Audio”, “Haptic Belt” and “Virtual Cane” methods with the “blind” sample are not normal distributed. For these methods, the following analysis does not apply.

According to the T-Test presented in the Table 8.11 the “Mixture” method is different between the “blind” and the “sight” sample. And according to the Figure 8.7, the mental demand of the “Mixture” method is statically higher in the “sight” sample than the “blind” sample.

According to the T-Test, both “Audio” and “Haptic Belt” are also different, but they are not normally distributed so it not possible this conclusion cannot be drawn. Also, the “Virtual Cane” is slightly higher in the “sight” sample than in the “blind”, but it was not detected in the test. All of these can be a consequence of a small sample size.



8.2.1.2 Analysis of the NASA-TLX score

The Table 8.14 presents the NASA score by each participant on each scene and they are plotted in the Figures 8.10 and 8.11. It is noticeable that after each “First” round the NASA score diminishes for both “sight” and “blind” participants.

TABLE 8.14 – NASA score felled by the participants.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Participant | Visual  Condition | Round |  |  |  |  |  |
| 001 | Sight | First | 7.83 | 10.17 | 9.83 | 7.00 | 9.000 |
|  |  | Return | 8.00 | 11.00 | 10.83 | 6.17 | 9.333 |
| 001C | Blind | First | 4.83 | 4.00 | 8.83 | 5.17 | 6.333 |
|  |  | Return | 4.17 | 4.00 | 6.67 | 4.50 | 6.167 |
| 002C | Blind | First | 6.33 | 4.83 | 4.83 | 9.00 | 7.000 |
|  |  | Return | 4.50 | 4.83 | 4.83 | 7.00 | 5.167 |
| 003 | Sight | First | 4.83 | 9.83 | 10.17 | 9.50 | 6.500 |
|  |  | Return | 4.33 | 6.67 | 9.67 | 7.83 | 4.833 |
| 003C | Blind | First | 4.00 | 4.00 | 5.33 | 6.67 | 3.500 |
|  |  | Return | 4.00 | 3.83 | 3.67 | 3.50 | 3.500 |
| 004 | Sight | First | 6.67 | 14.83 | 13.67 | 11.50 | 15.833 |
|  |  | Return | 6.83 | 11.83 | 11.83 | 10.83 | 12.167 |
| 004C | Blind | First | 9.83 | 10.00 | 12.67 | 9.67 | 11.000 |
|  |  | Return | 8.67 | 9.17 | 11.67 | 9.33 | 10.833 |
| 005 | Sight | First | 5.00 | 7.67 | 9.00 | 8.00 | 9.667 |
|  |  | Return | 5.00 | 7.67 | 8.67 | 7.67 | 6.000 |

The Figure 8.12 shows the NASA score between the rounds of each participant. This figure shows a noticeable difference between the two groups, meaning that probably the NASA score from the “sight” sample is higher than the one of the “blind” sample. This comparison can be also made with the data in the Table 8.15, that shows the average NASA score grouped by visual condition.

TABLE 8.15 – NASA-TLX score grouped by participant and visual Condition

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | 5.79 | 5.58 | 7.31 | 6.85 | 6.688 |
| Sight | 6.06 | 9.96 | 10.46 | 8.56 | 9.167 |

In the Figure 8.13 is plotted the average NASA score of each group and it also presents that the sighted participants felt a higher NASA score than the blinded participants.

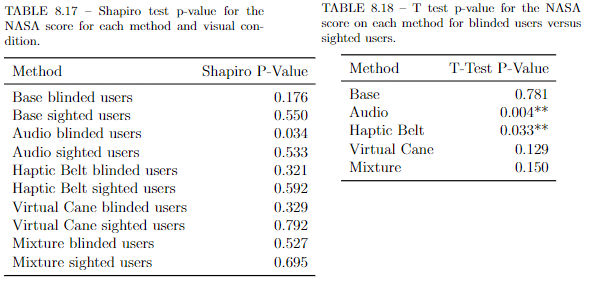
The Table 8.16 shows the NASA score variation grouped by visual condition and it also shows the difference between the mental demand of the “sight” sample and the ”blind” sample and how this score varies between the rounds.

TABLE 8.16 – NASA-TLX score grouped by participant and visual Condition.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | -13.7% | -3.1% | -15.9% | -21.5% | -7.6% |
| Sight | -1.4% | -11.1% | -3.0% | -9.9% | -20.8% |

The Figure 8.14 shows the variation of the NASA score of the “blind” sample and one can notice that the variation provoked on the ”Audio” method” is a lot lesser than the other ones.

The Shapiro–Wilk normality test on the Table 8.17 shows that these data are normally distributed. This means that further analysis can be applied for all of the methods.



According to the T-Test presented in the Table 8.18 the “Audio” and the “Haptic belt” caused a different NASA score between the “sight” sample and the ”blind” sample.

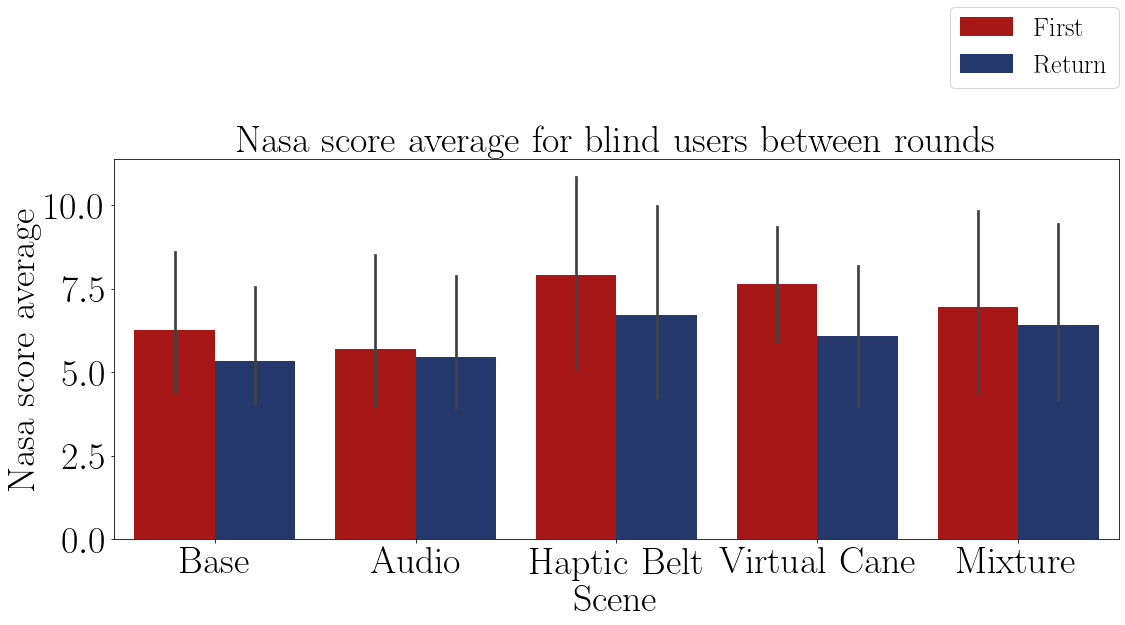


FIGURE 8.10 – Bar plot of the average Nasa-TLX score of the blind participants on each method.

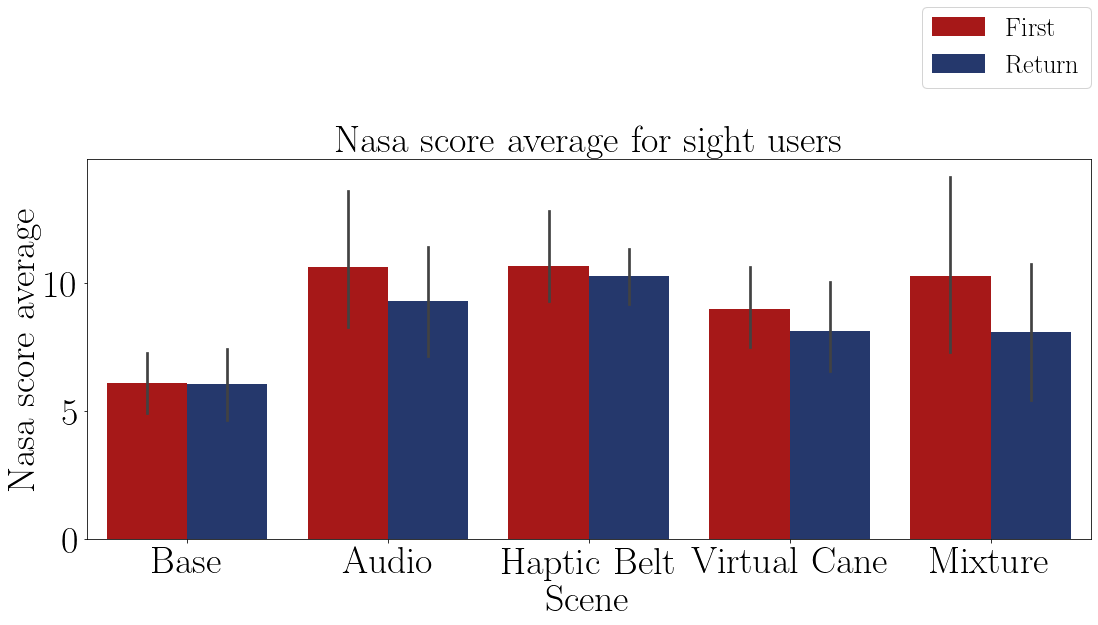


FIGURE 8.11 – Bar plot of the average Nasa-TLX score of the sighted participants on each method.

According to the Figure 8.12, and proved by the T-Test on Table 8.18, the “Audio” and “Haptic belt” caused a bigger NASA score when comparing both groups and analyzing the Figure 8.12 is noticeable that they are different. The other methods, excluding the “Base”, also seem to be different, but since the sample size was too small, maybe it was not able to detect the differences in them.

|  |  |
| --- | --- |
|  |  |
| FIGURE 8.12 – Boxplot of the average NASA-TLX score of the participants. | FIGURE 8.13 – Barplot of the average NASA score of each group. |

### 8.2.2 Adapted SAGAT

In this subsection, the SAGAT questionnaire is analyzed to assess the difference of perception between the two groups using the proposed assistive solutions. Its result may prove the necessity of a blind designer in order to increase the product effectiveness. As already explained before, for each question a participant could score 1 point or a fraction of it. The total score of each participant is presented on the Table 8.25 and they are plotted in the Figures 8.15 and 8.16. It is visually noticeable that both of the groups perform better the second time they visit the room.

TABLE 8.25 – Adapted SAGAT global score by participant and guidance method.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Participant | Visual  Condition | Round |  |  |  |  |  |
| 001 | Sight | First | 10.00 | 4.50 | 4.33 | 2.66 | 6.500 |
|  |  | Return | 10.00 | 6.00 | 5.00 | 5.00 | 4.500 |
| 001C | Blind | First | 6.25 | 5.50 | 5.33 | 5.83 | 3.500 |
|  |  | Return | 6.25 | 6.50 | 8.50 | 5.50 | 5.500 |
| 002C | Blind | First | 6.75 | 4.50 | 3.99 | 4.50 | 6.250 |
|  |  | Return | 5.25 | 5.00 | 4.00 | 6.50 | 8.500 |
| 003 | Sight | First | 10.00 | 6.75 | 5.99 | 3.99 | 6.750 |
|  |  | Return | 10.00 | 6.00 | 7.25 | 6.25 | 7.500 |
| 003C | Blind | First | 7.25 | 7.50 | 7.49 | 4.66 | 9.000 |
|  |  | Return | 10.00 | 10.00 | 8.50 | 9.00 | 9.000 |
| 004 | Sight | First | 10.00 | 7.25 | 7.99 | 5.99 | 8.250 |
|  |  | Return | 10.00 | 7.75 | 9.50 | 8.25 | 7.000 |
| 004C | Blind | First | 7.50 | 6.00 | 7.66 | 4.99 | 6.500 |
|  |  | Return | 9.00 | 6.00 | 9.25 | 7.25 | 9.000 |
| 005 | Sight | First | 10.00 | 3.00 | 3.16 | 3.99 | 4.000 |
|  |  | Return | 10.00 | 3.75 | 3.00 | 2.00 | 6.000 |

The Figure 8.17 shows the SAGAT score between the rounds of each participant and the performance of both groups was graphically similar.

In the Figure 8.18 the global average of both samples and it shows that the global average of the “sight” sample was lower than the “blind” sample.

The Tables 8.26 and 8.27 shows the average and the variation between the rounds of the SAGAT score grouped by visual condition, both of then present that the “blind” sample had a better performance than the “sight” sample.

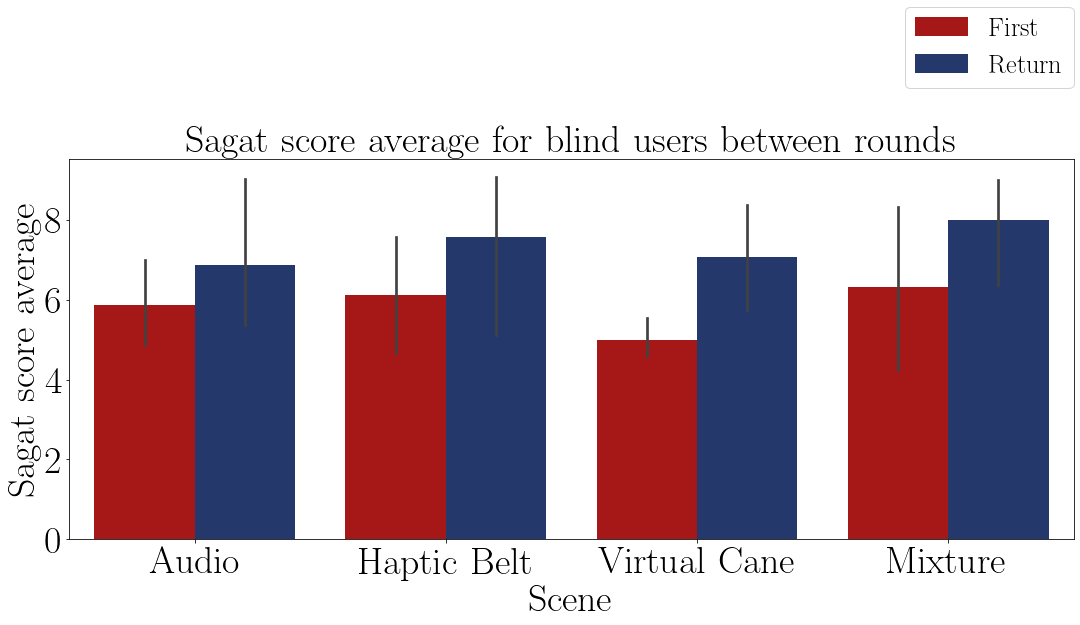


FIGURE 8.15 – Bar plot of the average Sagat score of the blind participants on each method.

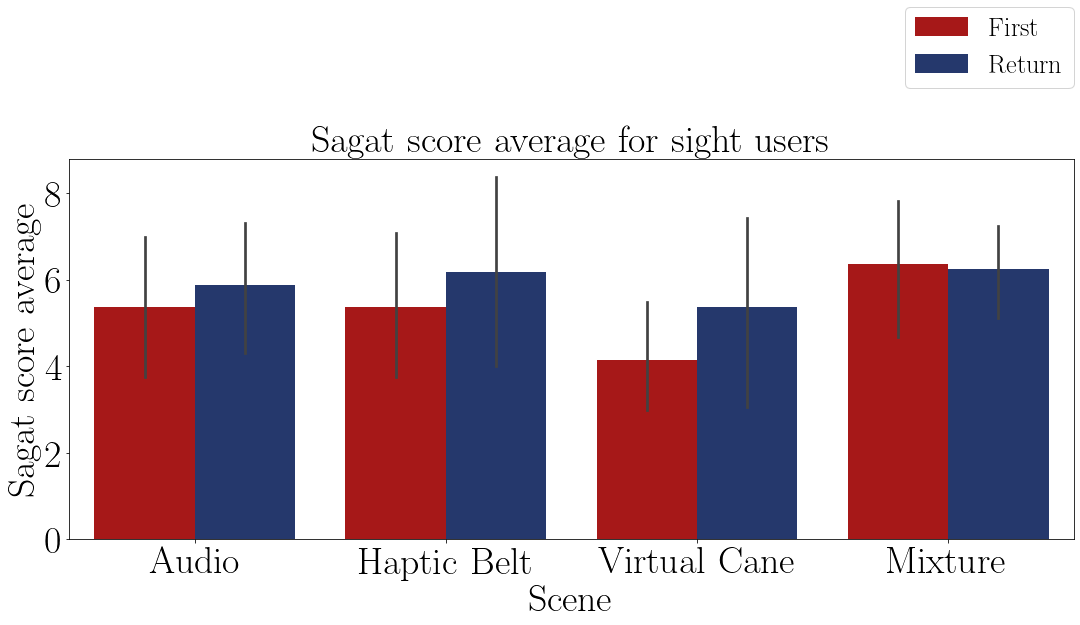


FIGURE 8.16 – Bar plot of the average Sagat score of the sighted participants on each method.

TABLE 8.26 – Adapted Sagat average global score grouped by participant and visual Condition.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | 7.28 | 6.38 | 6.84 | 6.03 | 7.156 |
| Sight | 10.00 | 5.62 | 5.78 | 4.77 | 6.312 |

The Shapiro–Wilk normality test on the Table 8.28 shows that these data are normally distributed, with a p-value higher than 0.05, then it is possible to perform a T-Test to guarantee that the ”blind” sample is different than the ”sight” sample.

According to the T-Test presented in the Table 8.29, the only method that showed

a difference in the Sagat score between the ”sight” sample and the ”blind” sample is the ”Base” method, which is expected. In the other methods both samples had a similar Sagat score.

TABLE 8.27 – Adapted Sagat global score variation grouped by participant and visual Condition

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Blind | 8.9% | 15.7% | 23.5% | 44.3% | 32.9% |
| Sight | 0.0% | 13.5% | 12.6% | 33.1% | 3.8% |

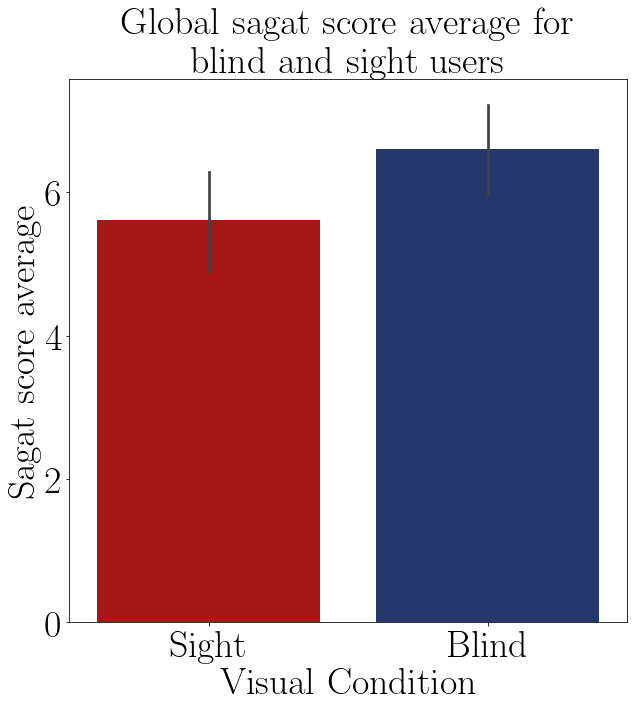
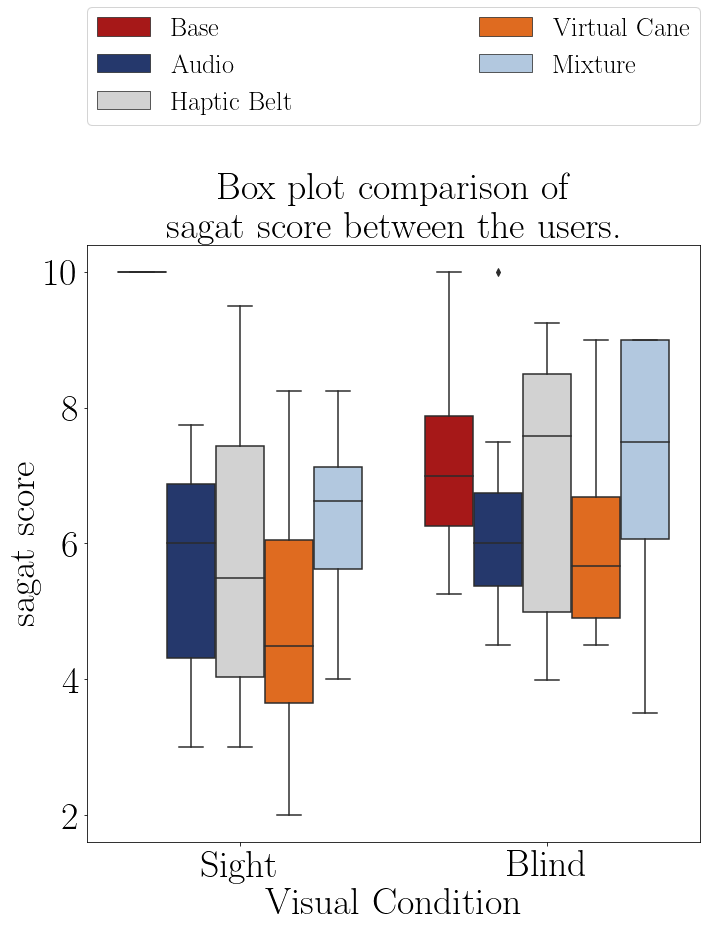
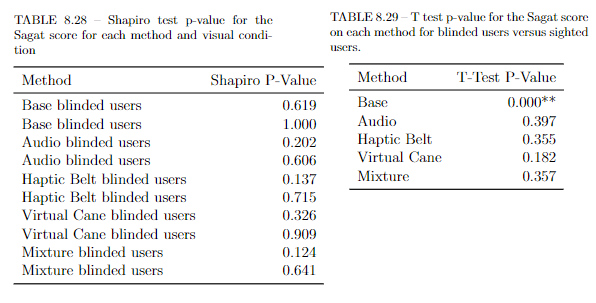


FIGURE 8.17 – Boxplot of the average SAGAT score of participants.

FIGURE 8.18 – Barplot of the average SAGAT score of each group.



The results from the Figure and from the T-Test showed that the both groups had a similar SAGAT score using the devices. (Questionario mal feito?)

## 8.2.2 Data from physiological sensors

The same sensors used for the first objective are used for the second objective. The expectations for all of the results is a difference between the “blind” sample and the “sight” sample.

### 8.1.2.1 Electrocardiogram (ECG) data

8.3.1.1 Analysis of the heartbeat frequency

The Table 8.32 presents the average heart rate by each participant on each scenes and they are plotted in the Figures 8.24 and their average plotted in the Figure 8.25. It possible to see that the distribution of the “blind” sample is rather different from the “sight” sample. The variation of the “blind” sample is bigger, which means that this sample had a greater mental workload relief than the “sight” sample.

TABLE 8.32 – ECG average BPM felled by the participants.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Baseline | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Part. | Visual  Condition | Round |  |  |  |  |  |  |
| 001 | Sight | First | 81.29 | 76.86 | 71.23 | 63.02 | 64.85 | 58.77 |
|  |  | Return |  | 72.88 | 73.18 | 61.18 | 66.78 | 66.26 |
| 001C | Blind | First | 78.33 | 75.75 | 60.71 | 71.17 | 59.07 | 68.24 |
|  |  | Return |  | 71.05 | 58.61 | 66.22 | 64.20 | 70.76 |
| 002C | Blind | First | 67.78 | 48.69 | 38.67 | 48.74 | 46.89 | 52.23 |
|  |  | Return |  | 52.46 | 47.58 | 58.97 | 56.75 | 58.25 |
| 003 | Sight | First | 77.38 | 74.98 | 63.47 | 71.80 | 70.90 | 72.76 |
|  |  | Return |  | 69.29 | 72.75 | 71.23 | 67.49 | 73.01 |
| 003C | Blind | First | 63.45 | 68.37 | 69.89 | 70.95 | 69.41 | 66.94 |
|  |  | Return |  | 67.34 | 67.44 | 69.68 | 68.82 | 67.37 |
| 004 | Sight | First | 65.32 | 72.97 | 66.85 | 62.45 | 65.94 | 67.86 |
|  |  | Return |  | 76.85 | 69.48 | 65.65 | 64.58 | 71.86 |
| 004C | Blind | First | 78.30 | 75.09 | 73.55 | 73.70 | 71.94 | 74.03 |
|  |  | Return |  | 74.74 | 74.79 | 74.02 | 72.69 | 67.34 |
| 005 | Sight | First | 71.25 | 70.18 | 71.34 | 66.93 | 66.46 | 67.06 |
|  |  | Return |  | 67.69 | 69.57 | 65.97 | 67.00 | 65.47 |

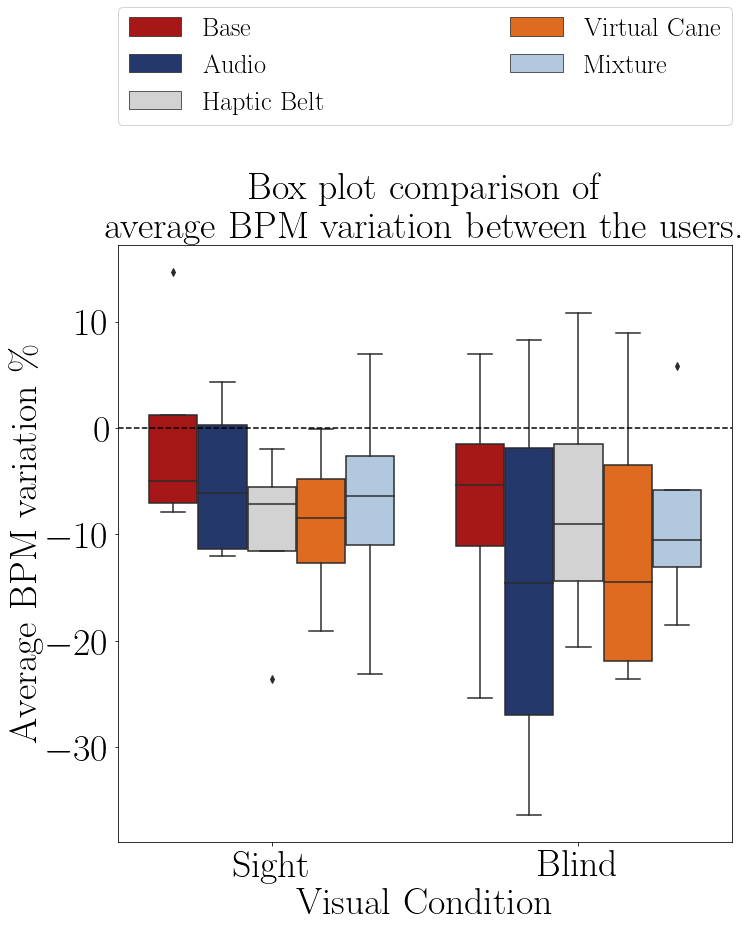


FIGURE 8.24 – Boxplot of the average heart rate of participants on each method.

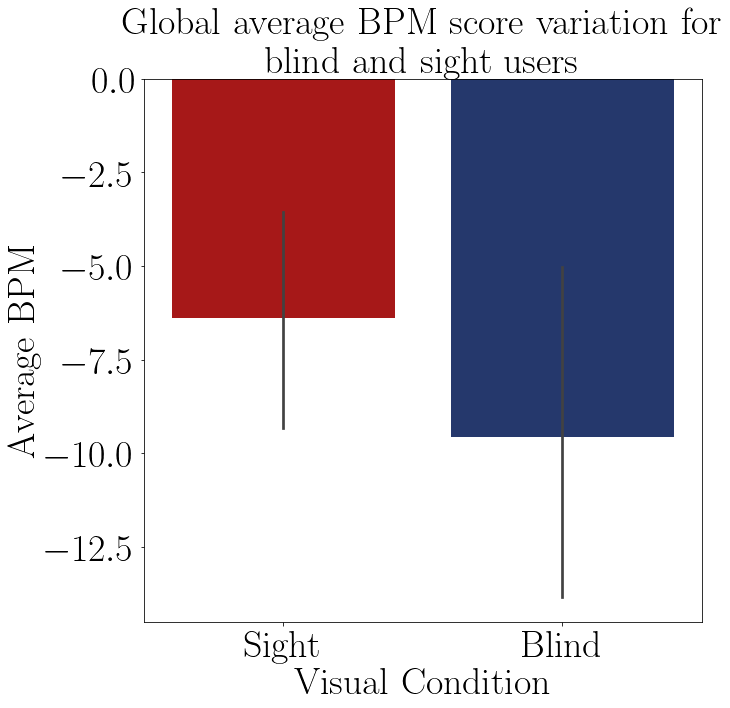


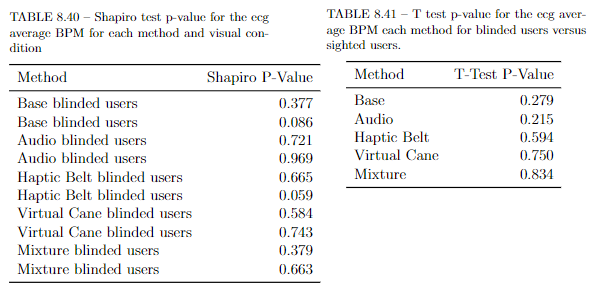
FIGURE 8.25 – Barplot of the average BPM score of each group.

The Table 8.33 show the average heartbeat frequency variation between the rounds of each group. In the presence of haptic devices, the sighted population had a similar relief than the “blind” population. This could by the fact of the sample profile, all students researcher of engineer.

TABLE 8.33 – ECG average BPM average in relation to the baseline grouped by participant and visual Condition.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic Belt | Virtual Cane | Mixture |
| Blind | -7.3% | -14.3% | -6.9% | -10.9% | -8.4% |
| Sight | -0.8% | -5.0% | -10.0% | -9.0% | -7.2% |

The Shapiro–Wilk normality test on the Table 8.40 shows that these data are normally distributed, with a p-value higher than 0.05, then it is possible to perform a T-Test to guarantee that the “blind” sample is different than the ”sight” sample. This test show that it is not possible to prove that there is a difference in one of the methods between the “blind” mental workload and the “sight” mental workload.



8.3.1.2 Analysis of the heartbeat variance

The Table 8.38 presents the standard deviation of the heartbeat interval by each participant on each scenes and their distribution and averages are plotted in the Figures 8.28 and 8.29 in that order. It possible to assume that both distribution and average are quite similar, in other words, both groups appear to had the same mental workload.

TABLE 8.38 – ECG Average SDNN felled by the participants.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Baseline | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Part. | Visual  Condition | Round |  |  |  |  |  |  |
| 001 | Sight | First | 37.52 | 82.73 | 82.19 | 134.53 | 134.77 | 225.41 |
|  |  | Return |  | 84.96 | 69.48 | 318.75 | 116.00 | 136.51 |
| 001C | Blind | First | 78.55 | 81.29 | 107.06 | 124.74 | 163.97 | 129.05 |
|  |  | Return |  | 120.72 | 130.88 | 131.59 | 157.59 | 124.79 |
| 002C | Blind | First | 93.77 | 73.76 | 98.86 | 81.14 | 33.98 | 79.29 |
|  |  | Return |  | 108.94 | 49.63 | 42.81 | 114.06 | 107.55 |
| 003 | Sight | First | 45.40 | 58.07 | 79.60 | 51.78 | 68.68 | 60.84 |
|  |  | Return |  | 21.30 | 45.71 | 40.93 | 66.32 | 47.82 |
| 003C | Blind | First | 26.14 | 36.87 | 38.32 | 35.10 | 42.39 | 43.69 |
|  |  | Return |  | 52.75 | 41.20 | 44.26 | 42.60 | 46.14 |
| 004 | Sight | First | 91.79 | 120.51 | 121.13 | 154.72 | 128.48 | 125.95 |
|  |  | Return |  | 139.86 | 100.37 | 122.56 | 140.12 | 119.26 |
| 004C | Blind | First | 20.98 | 70.73 | 86.83 | 62.56 | 85.90 | 70.47 |
|  |  | Return |  | 71.95 | 74.89 | 70.02 | 66.09 | 104.04 |
| 005 | Sight | First | 80.61 | 44.50 | 87.69 | 120.52 | 88.59 | 102.80 |
|  |  | Return |  | 59.77 | 93.21 | 122.84 | 141.31 | 96.03 |

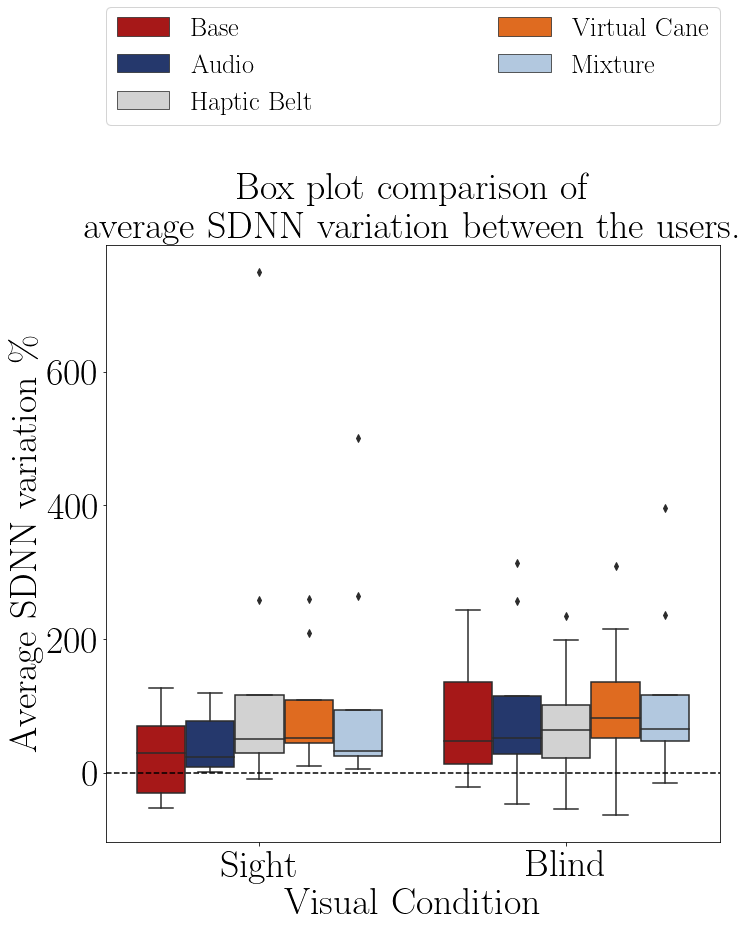


FIGURE 8.28 – Boxplot of the average heart rate of the participants on each method.

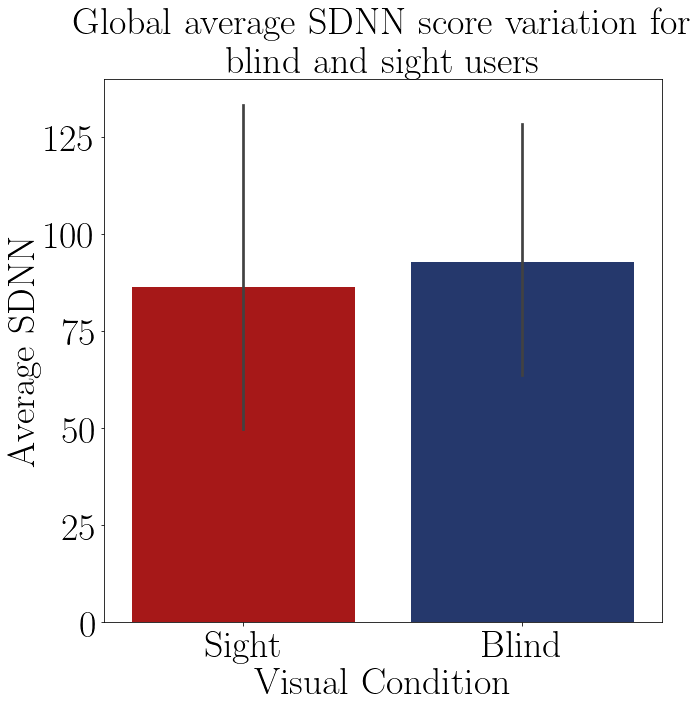


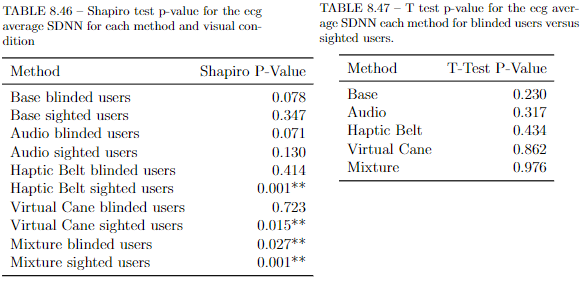
FIGURE 8.29 – Barplot of the average SDNN score of each group.

The Table 8.39 shows the variation of the heartbeat in each round of each group. Its conclusion is quite similar to the analogue table for the heartbeat analysis, that the presence of the haptic devices appears to be help the sighted user more than the blinded users.

TABLE 8.39 – ECG Average SDNN average in relation to the baseline grouped by participant and visual Condition.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual Condition | Base | Audio | Haptic Belt | Virtual Cane | Mixture |
| Blind | 84.4% | 92.1% | 74.3% | 102.1% | 112.2% |
| Sight | 29.3% | 43.2% | 152.0% | 92.9% | 114.7% |

The Shapiro–Wilk normality test on the Table 8.46 shows that some of the methods are not normally distributed. All of the methods with the presence of haptic devices for the sighted user are not normally distributed, and the “Mixture” method for the blind users is not normally distributed. That means that the T-Test analyze cannot be made those methods. But according to the T-Test, it is not possible to assume that all of the methods are different between the groups. That means the it is not possible to prove that the mental workload of the “blind” sample is different than the mental workload of the “sight” sample.



### 8.2.2.2 Galvanic skin reaction.

The GSR analysis is also made by analyzing its average and the accumulated value and comparing both features between both blind and sample groups. As mentioned before, there was no influence of the temperature and the GSR sensor was worn on the left hand for right-handed participant and on the right hand for left-handed participants.

8.3.2.1 Analysis of the GSR average

The Table 8.44 presents the average skin conductance by each participant on each scene and their distribution is plotted in the Figures 8.32 and the groups averages in the Figure 8.33. It is possible to see that only for the “Audio” method there is significant difference, and that the averages are rather similar.

TABLE 8.44 – Average GSR felled by the participants.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Baseline | Base | Audio | Haptic  Belt | Virtual  Cane | Mixture |
| Part. | Visual  Condition | Round |  |  |  |  |  |  |
| 001 | Sight | First | 4.27 | 8.80 | 15.19 | 15.67 | 15.19 | 14.15 |
|  |  | Return |  | 11.48 | 14.95 | 15.09 | 15.72 | 21.52 |
| 001C | Blind | First | 0.37 | 0.48 | 1.03 | 3.14 | 3.79 | 3.90 |
|  |  | Return |  | 0.83 | 1.58 | 2.81 | 4.04 | 4.57 |
| 002C | Blind | First | 0.17 | 0.91 | 0.23 | 0.17 | 0.17 | 0.17 |
|  |  | Return |  | 0.43 | 0.17 | 0.16 | 0.17 | 0.17 |
| 003 | Sight | First | 0.19 | 0.19 | 0.17 | 0.17 | 0.17 | 0.17 |
|  |  | Return |  | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 |
| 003C | Blind | First | 0.30 | 0.56 | 0.56 | 0.62 | 0.85 | 1.09 |
|  |  | Return |  | 0.62 | 0.63 | 0.65 | 0.92 | 1.06 |
| 004 | Sight | First | 2.60 | 9.71 | 11.18 | 12.60 | 12.92 | 10.34 |
|  |  | Return |  | 10.89 | 11.97 | 12.25 | 13.47 | 10.16 |
| 004C | Blind | First | 1.24 | 2.34 | 3.07 | 3.49 | 2.28 | 2.23 |
|  |  | Return |  | 2.57 | 2.95 | 3.20 | 2.21 | 2.24 |
| 005 | Sight | First | 0.47 | 1.88 | 1.58 | 1.44 | 1.37 | 1.33 |
|  |  | Return |  | 1.66 | 1.53 | 1.47 | 1.49 | 1.33 |

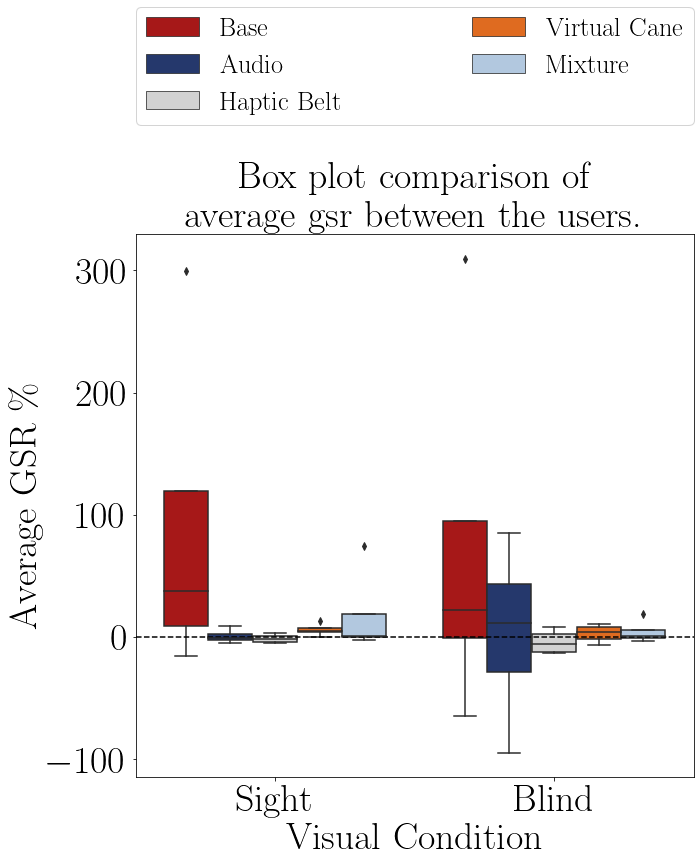


FIGURE 8.32 – Boxplot of the average skin conductace of the participants on each method.

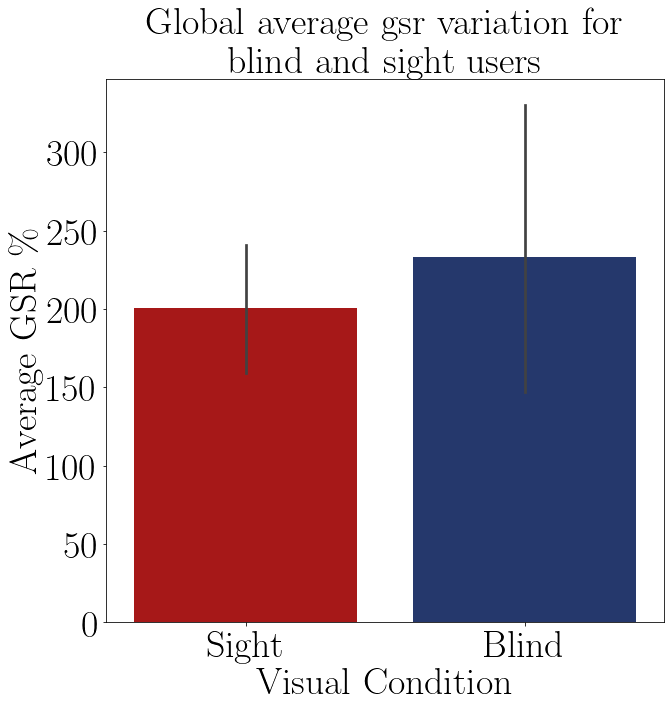
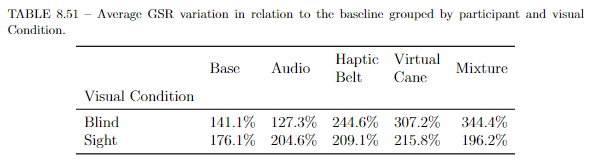
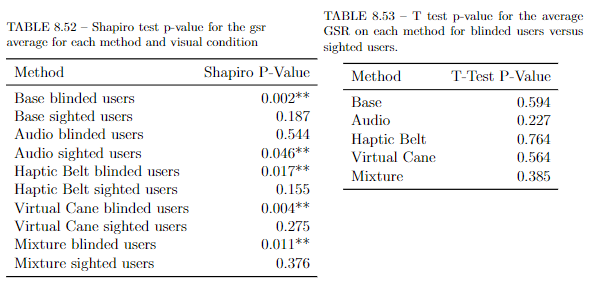


FIGURE 8.33 – Bar plot of the average GSR of of each group.

The Table 8.51 presents the average skin conductance by groups on each scene It is possible to see that the averages are rather different between the groups.

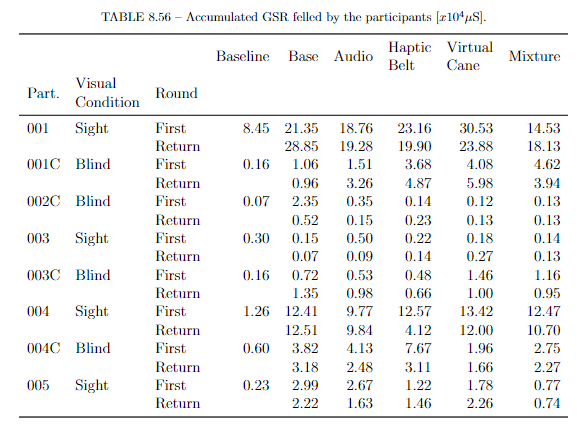


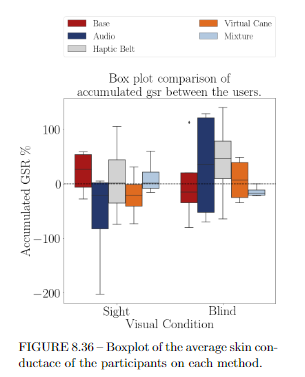
The Shapiro–Wilk normality test on the Table 8.52 shows that there are one group in each method that does not have a normal distribution, so is not possible to prove that they are different

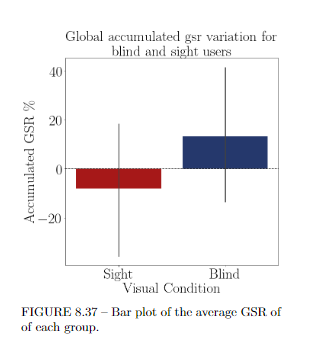


8.3.2.2 Analysis of the accumulated GSR

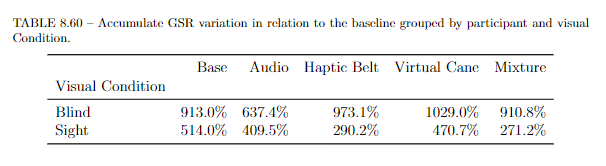
The Table 8.56 presents the accumulated skin conductance by each participant on each scene and their distribution is plotted in the Figures 8.36 and the groups averages in the Figure 8.37. It is possible to see that only for the “Audio” method there is a significant difference, and that the averages are quite different, with the sighted participants average being smaller than the “blind” average, meaning that the sighted participants felt a minor mental workload or arousal.



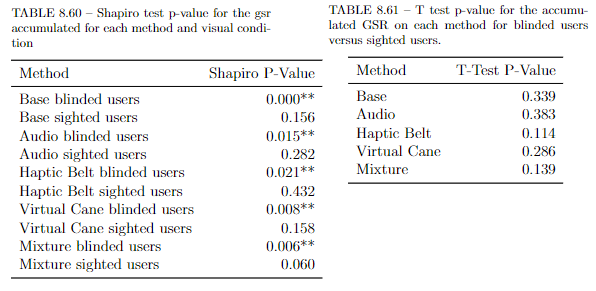




The Table 8.60 presents the average accumulated skin conductance by groups on each scene It is possible to see that the averages are rather different between the groups and that the sighted sample has lower accumulates skin conductance.



The Shapiro–Wilk normality test on the Table 8.52 shows that all of the “blind” sample methods are not normally distributed, so it is not possible to assume that there is a statistical difference between the groups.



## 2.4 Final remarks

The NASA-TLX questionnaire show that there was a difference about the mental workload between the sample groups and the box plot Figures shown that the difference is in all the methods, but there were some tests they could only prove some of the differences, and is this is due the small sample size. But the adapted SAGAT was not able to prove any difference and there are two reasons for it:

* Bad adaptation or a poor design;

Even though the questionnaire was made with the assistance of a blind user, it still an adaptation and it is the first time it was tested, so is just normal that it does not work as expected.

* Sample group profiling.

As already explained before, the profile of the “blind” sample group was very wide and that can impact negatively in their performance. But the opposite effect may had happened with the “sight” sample group, that were integrated basically with researchers and students of engineer, a group that is typically involved with technological devices, and with a age range from 22 to 31 with an average of 27.5 years. This could have selected users that would perform better when using the HMD and being able to feel present inside a virtual environment.

The ECG sensors shown only a slight difference in the boxplot figures and the T-Test did not prove any difference. This probably happened for similar reasons as pointed objective above: The noise caused by the users’ movement and a small sample size. The GSR showed some small differences in the boxplot, but no method had both groups normally distributed, main condition to use the T Test. Again, this could be because of the users movement and small sample size.

Besides these results, the “sighted” sample also commented the experiment. They all felt a lot more insecurity when walking, exploring and even when hand guided by the researcher before the start of the round. The “blind” sample group was already used to bumping their body when exploring new closed quarters. The “sighted” group did not want that to happen and approached the furniture with a lot more caution. They also noticed the lack of precision of the haptic devices, but they did rely more on then to navigate.