1 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.

1.1 Evaluation of assistive device from a human factors' perspective in a virtual environment

1.1.1 Subjective data

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.

1.1.1.1 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.

1.1.1.1.1 Analysis of the mental demand scale

The Table 1.1 presents the mental demand averages by each blinded participant on each scene and they are plotted in the Figures 1.1. The Figure 1.1 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.

TABLE 1.1 – Mental demand felled by the blinded participants.

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

The Figure 1.2 presents a box plot with the mental demand grouped by method. This Figure shows that there may be two different groups, one with lower demand formed by the "Base" and the "Audio" method, and another with the higher demand. The Figure 1.3 presents a box plot with the mental demand grouped by the rounds. This figure shows that both rounds have similar variations.

The Table 1.2 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.

The Figures 1.4 and 1.5 shows the distribution and variance of the Table 1.1. These Figures shows that the data are normally distributed and that the methods have a similar

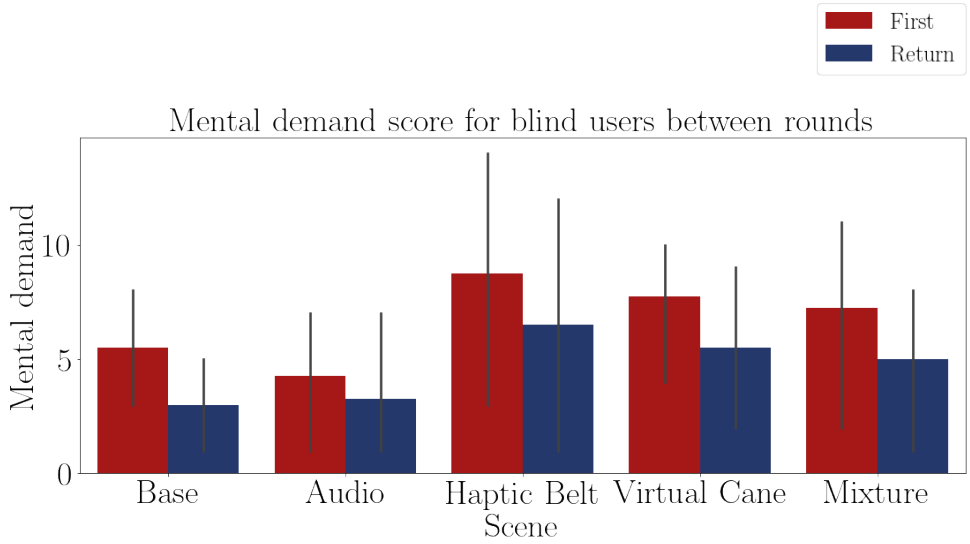

FIGURE 1.1 – Barplot of the average mental demand of the blind participants on each method.

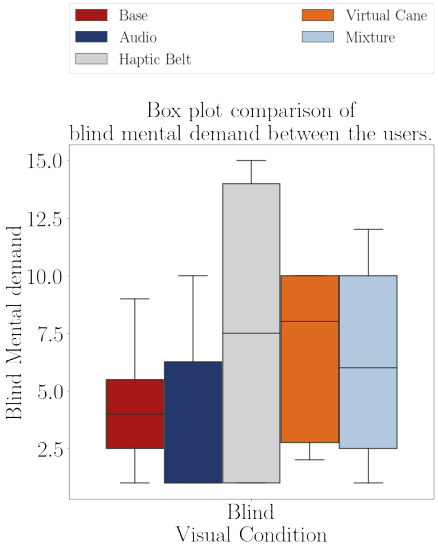

FIGURE 1.2 – Boxplot of the mental demand of the blind participants grouped by method.

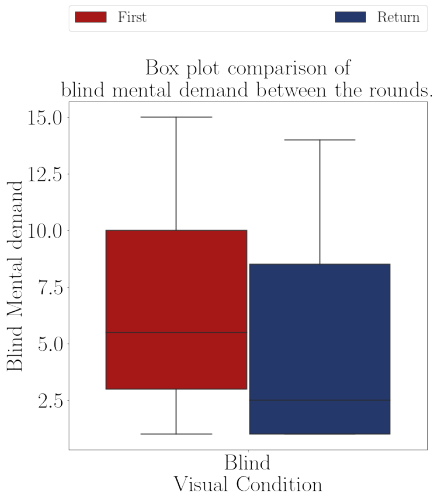

FIGURE 1.3 – Boxplot of the mental demand of the blind participants grouped by round.

TABLE 1.2 – 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

variance. The Table 1.3 shows the Anova test p-values of the mental demand of the "blind" sample between the guidance methods. The method's and the round's p-values indicates that there is no influence from them in the mental demand. The interaction between the methods and the round also does not influences the mental demand.

The Table 1.4 shows the average of the mental demand variation between the rounds.

TABLE 1.3 – Anova p-value for the mental demand average on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (Blocks)	298.475	3	99.492	8.133	
Methods	85.150	4	21.288	1.740	0.170
Rounds	42.025	1	42.025	3.436	0.075
Interaction	2.850	4	0.712	0.058	0.993
Experimental Error	330.275	27	12.232		
Total	758.775	39			

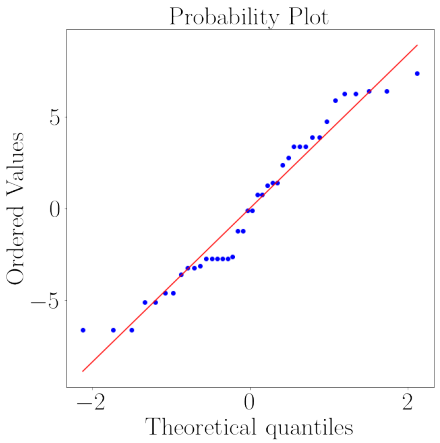

FIGURE 1.4 – QQ plot of the mental demand of the blind participants on each method.

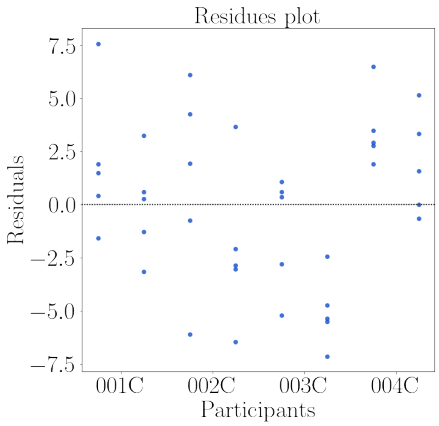

FIGURE 1.5 – Residual plot of the mental demand score the blind participants on each method.

This table shows that the mental demand variation from the "Audio" has the lower variation, and the rest are similar variations.

TABLE 1.4 – Mental demand variation grouped by participant and visual condition

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	-2.5	-1.0	-2.2	-2.2	-2.2

The Figures 1.6 and 1.7 shows the distribution and variance of the mental demand variation of the Table 1.1. These Figures shows that the data are normally distributed and that the methods have a similar variance. The Table 1.5 shows the Anova test p-value of the mental demand of the "blind" sample between the guidance methods. The p-value indicates that there is no influence of the methods in the variation of mental demand between the rounds.

To close up, according to the ANOVA test at Table 1.3 there is no influence in the tested methods in the participants mental demand, but at the Figure 1.2 it is posible to notice that there is at least two different groups of mental demand reactions, one formed by

TABLE 1.5 – Anova p-value for the mental demand variation on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (blocks)	15.750	3	1.425	0.674	
Method	5.700	4	5.250	0.183	0.943
Experimental error	93.500	12	7.792		
Total	114.950	19			

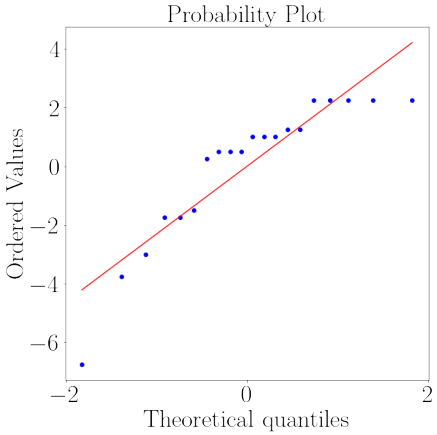

FIGURE 1.6 – Residual plot of the mental demand variation of the blind participants on each method.

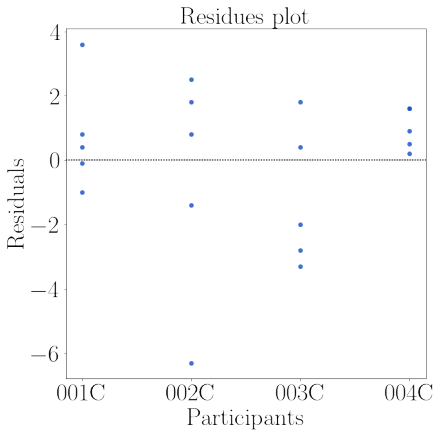

FIGURE 1.7 – Residual plot of the mental demand variation of the sighted participants on each method.

the "Base" and the "Audio" methods and another formed by the rest of the methods. The first group has lower mental demand than the last. That could mean that the presence of a haptic device increases the mental demand of the navigation activity for the BVI users. This was not reflected in the ANOVA results because of the small sample size.

1.1.1.1.2 Analysis of the NASA-TLX score

The Table 1.6 presents the NASA-TLX score averages by each blinded participant on each scene and they are plotted in the Figures 1.8. The Figure 1.8 shows a similar behaviour of the mental demand barplot at Figure 1.1, all NASA-TLX score decreased from the "First" to the "Return" round. This a kind of learning between the rounds.

The Figure 1.9 presents a box plot with the NASA-TLX score grouped by method. This Figure shows it is possible to split the methods in two different groups, one with lower demand formed by the "Base" and the "Audio" method, and another with the higher demand, similar as it was with the mental demand in the 1.2. It appears that the presence of the an haptic device elevated the NASA-TLX score. The Figure 1.10 presents a box plot with the NASA-TLX score grouped by the rounds. This figure shows that both rounds have similar variations.

TABLE 1.6 – NASA-TLX score felled by the blinded participants.

Participant	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001C	First	4.833	4.000	8.833	5.167	6.333
	Return	4.167	4.000	6.667	4.500	6.167
002C	First	6.333	4.833	4.833	9.000	7.000
	Return	4.500	4.833	4.833	7.000	5.167
003C	First	4.000	4.000	5.333	6.667	3.500
	Return	4.000	3.833	3.667	3.500	3.500
004C	First	9.833	10.000	12.667	9.667	11.000
	Return	8.667	9.167	11.667	9.333	10.833

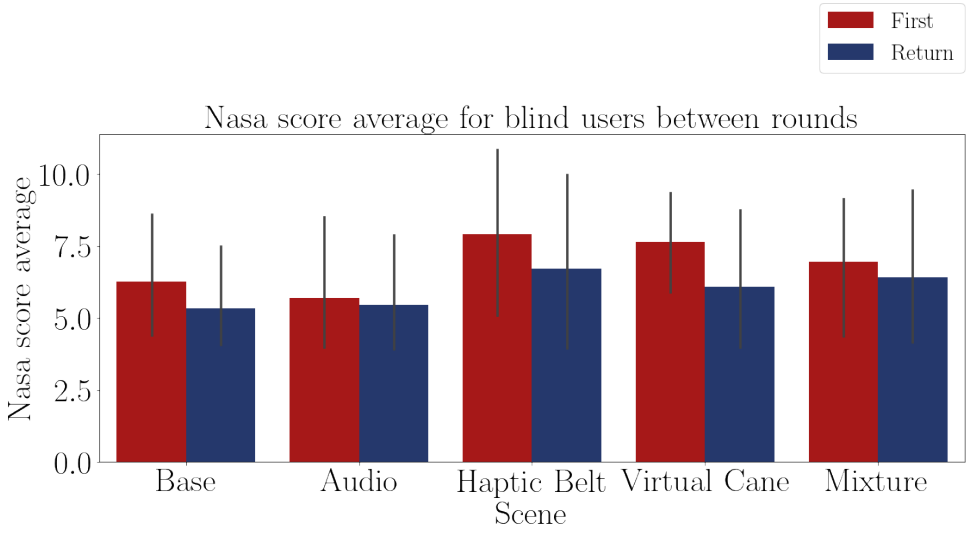

FIGURE 1.8 – Barplot of the average NASA-TLX score of the blind participants on each method.

The Table 1.7 shows the average NASA-TLX score in the "blind" sample and is possible to notice how the average score by the "blind" sample was lower during the "Audio" and the "Base" methods.

TABLE 1.7 – NASA-TLX average 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 Figures 1.11 and 1.12 shows the distribution and variance of the Table 1.6. These Figures shows that the data are normally distributed and that the methods have a similar variance. The Table 1.8 shows the Anova test p-value of the NASA-TLX score of the "blind" sample between the guidance methods. The p-values indicates that some methods

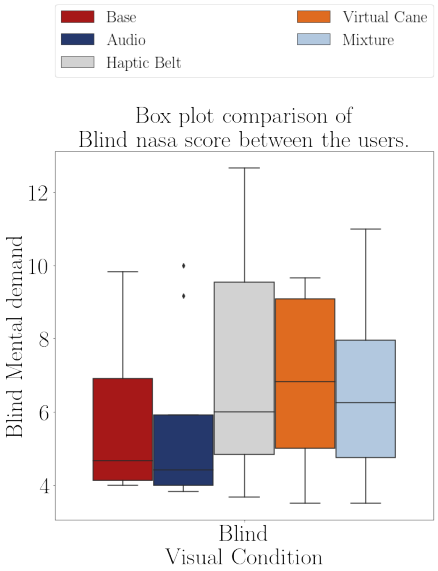

FIGURE 1.9 – QQ plot of the NASA-TLX score of the blind participants on each method.

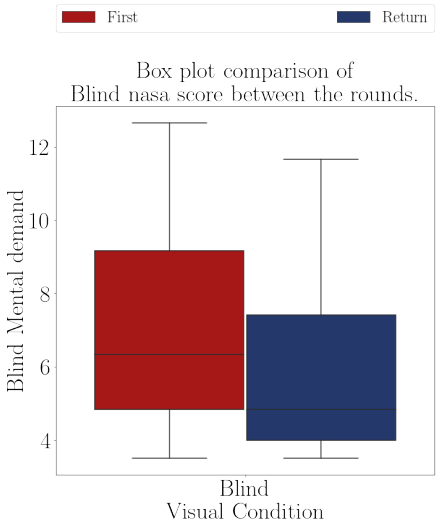

FIGURE 1.10 – Residual plot of the NASA-TLX score the blind participants on each method.

have influence on the NASA-TLX score and that the rounds also influences the score. On the other way, their interaction, has no influence on the score.

TABLE 1.8 – Anova p-value for the mental demand average on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (Blocks)	211.041	3	70.347	51.869	
Methods	17.185	4	4.296	3.168	0.029**
Rounds	7.951	1	7.951	5.862	0.022**
Interaction	2.115	4	0.529	0.390	0.814
Experimental Error	36.619	27	1.356		
Total	274.910	39			

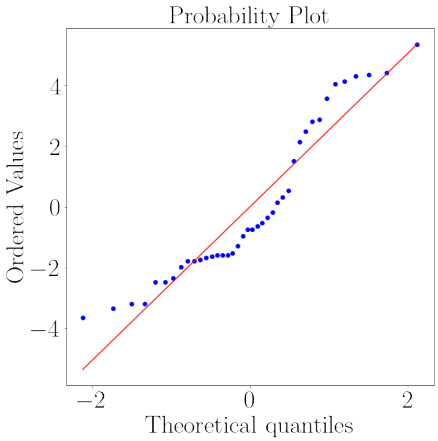

FIGURE 1.11 – QQ plot of the NASA-TLX score variation of the blind participants on each method.

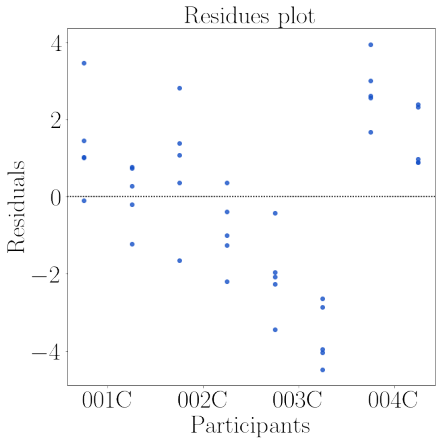

FIGURE 1.12 – Residual plot of the NASA-TLX score variation the blind participants on each method.

The Table 1.9 presents the conclusion of a pairwise Fisher LSD test of the blind NASA-TLX score between all the guidance methods. The results show that only the "Audio" has a similar NASA-TLX score as the "Base" method, as it was also posible to notice at Figure 1.9.

TABLE 1.9 – Cross validation p-value for the mental demand average on each method for blinded users.

Method			Analysis
Base	X	Audio	$H_0 : \mu_{Base} = \mu_{Audio}$
Base	X	Haptic Belt	$H_1 : \mu_{Base} \neq \mu_{HapticBelt} **$
Base	X	Virtual Cane	$H_1 : \mu_{Base} \neq \mu_{VirtualCane} **$
Base	X	Mixture	$H_1 : \mu_{Base} \neq \mu_{Mixture} **$
Audio	X	Haptic Belt	$H_1 : \mu_{Audio} \neq \mu_{HapticBelt} **$
Audio	X	Virtual Cane	$H_1 : \mu_{Audio} \neq \mu_{VirtualCane} **$
Audio	X	Mixture	$H_1 : \mu_{Audio} \neq \mu_{Mixture} **$
Haptic Belt	X	Virtual Cane	$H_1 : \mu_{HapticBelt} \neq \mu_{VirtualCane} **$
Haptic Belt	X	Mixture	$H_1 : \mu_{HapticBelt} \neq \mu_{Mixture} **$
Virtual Cane	X	Mixture	$H_0 : \mu_{VirtualCane} = \mu_{Mixture}$

The Table 1.10 shows the average of the NASA-TLX score variation between the rounds. This table shows that the variation from the "Audio" was the lowest variation and the highest variation was the "Virtual Cane".

TABLE 1.10 – NASA-TLX score grouped by participant and visual Condition.

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	-0.92	-0.25	-1.21	-1.54	-0.54

The Figures 1.13 and 1.14 shows the distribution and variance of the NASA-TLX score variation of the Table 1.6. These Figures shows that the data are normally distributed and that the methods have a similar variance. The Table 1.11 shows the Anova test p-value of the NASA-TLX score of the "blind" sample between the guidance methods. The p-value indicates that there are no difference between the variation of any method.

TABLE 1.11 – Anova p-value for the NASA score variation on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (blocks)	0.660	3	1.058	0.228	
Method	4.231	4	0.220	1.097	0.402
Experimental error	11.569	12	0.964		
Total	16.460	19			

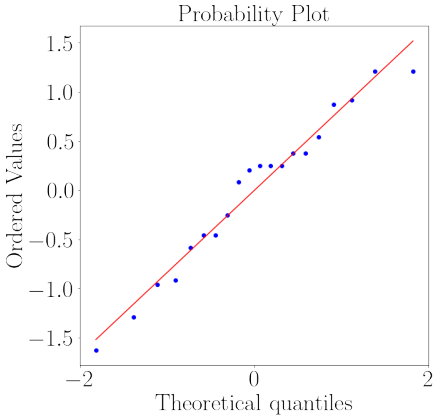

FIGURE 1.13 – Bar plot of the average NASA-TLX score of the blind participants on each method.

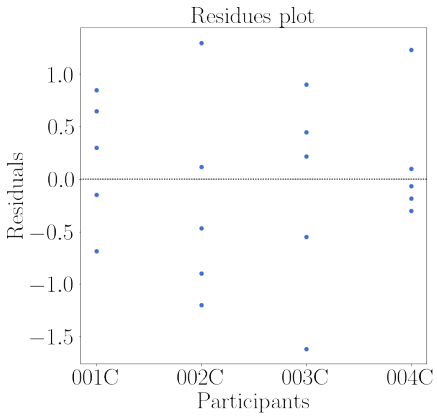

FIGURE 1.14 – Bar plot of the average NASA-TLX score of the sighted participants on each method.

To close up, according to the LSD test at Table 1.9 only the "Audio" method has a NASA-TLX score that could be said to be similar to the "Base" method, which indicates that the existance of an haptic device increased the NASA-TLX score. The 1.8 also concludes that the rounds and the interaction between the rounds and the methods have no influence on the NASA-TLX score.

1.1.1.2 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 1.12 and they are plotted in the Figures 1.15, where it is visually noticeable that the performance better the second time they visit the room.

TABLE 1.12 – SAGAT global score felled by the blinded participants.

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

The boxplot in the Figure 1.16 shows that there are two groups of scores one with the "Base", "Haptic Belt" and the "Mixture" methods, and the second group with the

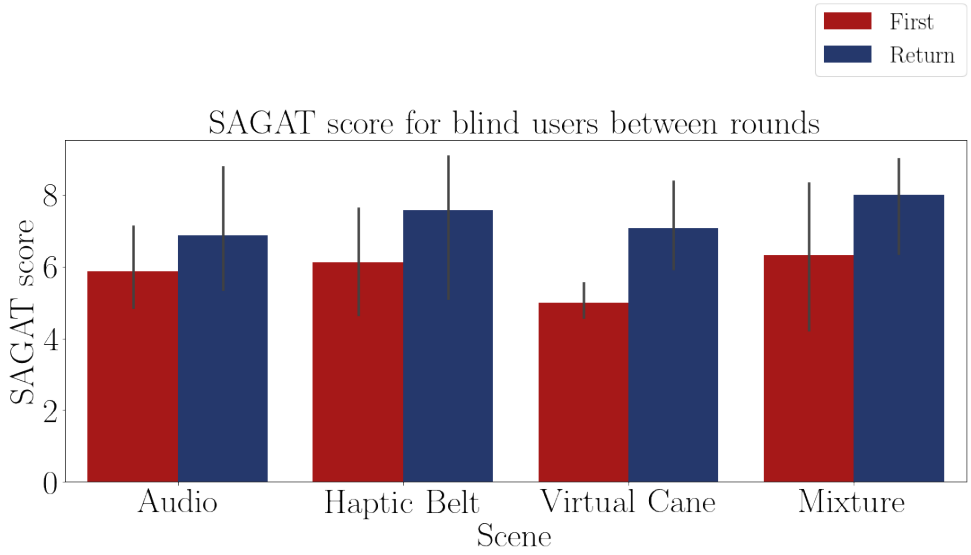

FIGURE 1.15 – Barplot of the average SAGAT score of the blind participants on each method.

"Audio" and the "Virtual Cane" methods. The first group scored higher than the second one. The Figure 1.17 shows a noticible difference between the scores when grouped by their corresponding round.

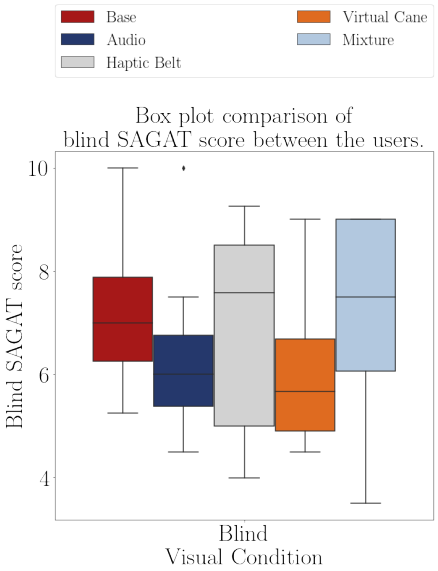

FIGURE 1.16 – Boxplot of the SAGAT score of the blind participants grouped by method.

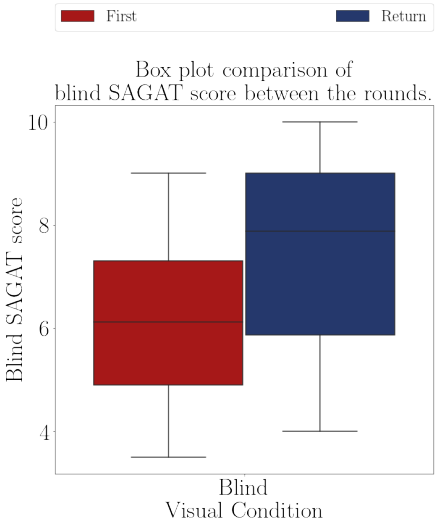

FIGURE 1.17 – Boxplot of the SAGAT score of the blind participants grouped by round.

The Table 1.13 shows the average SAGAT score in the "blind" sample and is possible to notice how the average score by the "blind" sample was lower during the "Audio" and the "Base" methods.

The Figures 1.24 and 1.19 shows the distribution and variance of the Table 1.12. These Figures shows that the data are normally distributed and that the methods have a similar variance. The Table 1.14 shows the Anova test p-value of the SAGAT score of the "blind" sample. The round's p-values indicates that some have influence on the SAGAT score.

TABLE 1.13 – 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

Meaning that the participants did learn information about the room between the "First" and "Return" round. The method and the interaction between it and the round has no influence on the SAGAT score.

TABLE 1.14 – Anova p-value for the SAGAT score on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (Blocks)	48.231	3	16.077	9.731	
Methods	8.922	4	2.230	1.350	0.277
Rounds	18.975	1	18.975	11.485	0.002**
Interaction	2.391	4	0.598	0.362	0.834
Experimental Error	44.608	27	1.652		
Total	123.127	39			

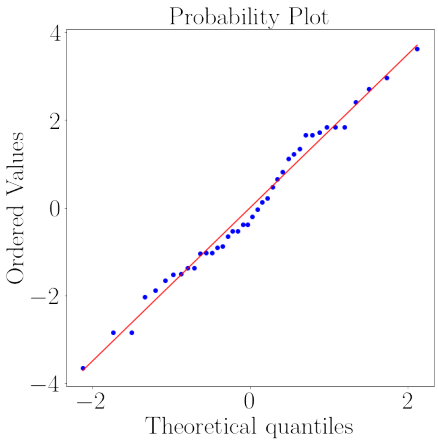

FIGURE 1.18 – QQ plot of the SAGAT score of the blind participants on each method.

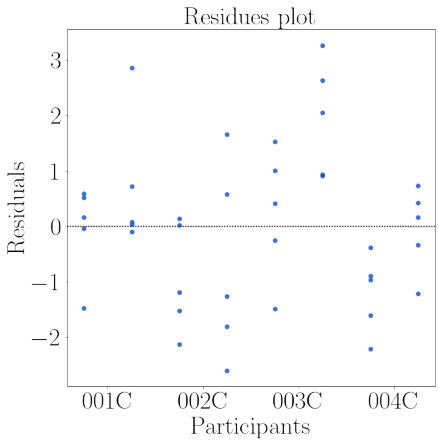

FIGURE 1.19 – Residual plot of the SAGAT score the blind participants on each method.

The Table 1.15 shows the average of the SAGAT score variation between the rounds. This table shows that the variation from the "Base" and the "Audio" was the lowest variation and the highest variation was the "Virtual Cane".

The Figures 1.20 and 1.21 shows the distribution and variance of the SAGAT score variation of the Table 1.12. These Figures shows that the data are normally distributed and that the methods have a similar variance. The Table 1.11 shows the Anova test p-value of the SAGAT score of the "blind" sample between the guidance methods. The p-value indicates that there are no difference between the variation in any method.

TABLE 1.15 – Adapted Sagat global score variation grouped by participant and visual Condition

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	8.93	15.66	23.49	44.30	32.90

TABLE 1.16 – Anova p-value for the Sagat score variation on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (blocks)	1176.902	3	782.885	0.473	
Method	3131.542	4	392.301	0.944	0.472
Experimental error	9956.458	12	829.705		
Total	14264.902	19			

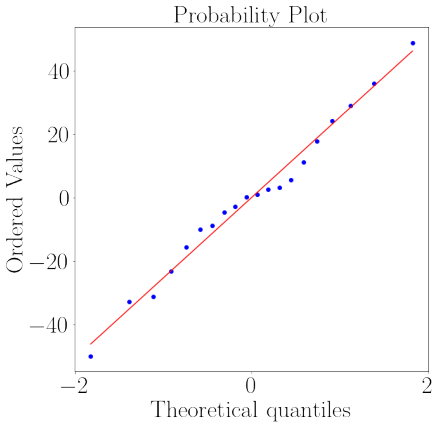

FIGURE 1.20 – QQ plot of the SAGAT score variation of the blind participants on each method.

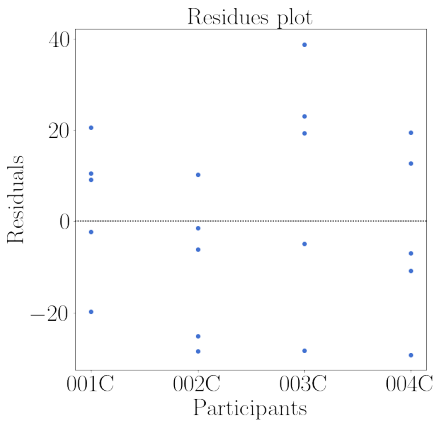

FIGURE 1.21 – Residual plot of the SAGAT score variation of the blind participants on each method.

To close up, according to the ANOVA test at Table 1.14 the methods caused no reaction on the SAGAT score, but the rounds did. That means that the participants were able in all methods to learn a little about their environment and that learning impacted their environmental perception in the next round. The fact that the test has not found any influence of the methods on the SAGAT score may be because of the small sample size, since it is posible to notice a difference between the methods at Figure 1.16. Also the interaction between method and round caused no influence in the Sagat score. According to the ANOVA test at Table 1.16, the methods did not influenced the SAGAT score.

1.1.1.3 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 1.17 shows the score of each method and they are plotted in the Figure 1.22. The Figure show a disatisfaction with the haptic devices alone.

TABLE 1.17 – Guidance method questionnaire score felled by the blinded participants.

Participant	Audio	Haptic Belt	Virtual Cane	Mixture
001C	0.774	0.543	0.629	0.865
002C	0.857	0.743	0.543	0.935
003C	0.929	0.571	0.543	0.745
004C	0.881	0.486	0.400	0.730

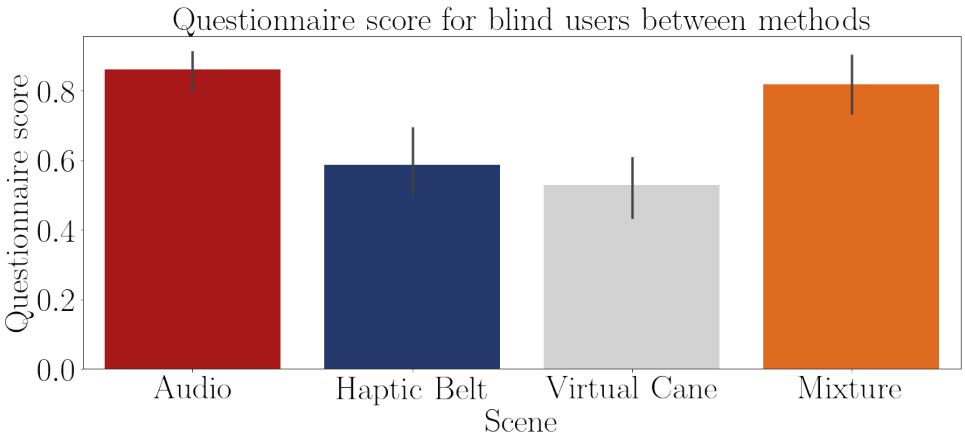

FIGURE 1.22 – Barplot of the average questionaire score of the blind participants on each method.

The Table 1.18 show the the average questionnaire score on each method. It also shows a disatisfaction with the haptic devices alone.

TABLE 1.18 – 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 Figures 1.24 and 1.19 shows the distribution and variance of the Table 1.12. These Figures shows that the data are normally distributed and that the methods have a similar variance. The Table 1.14 shows the Anova test p-value of the SAGAT score of the "blind" sample. The p-values indicates that the method have influence on the questionnaire score. Meaning that the participants had differents level os satisfaction about each method.

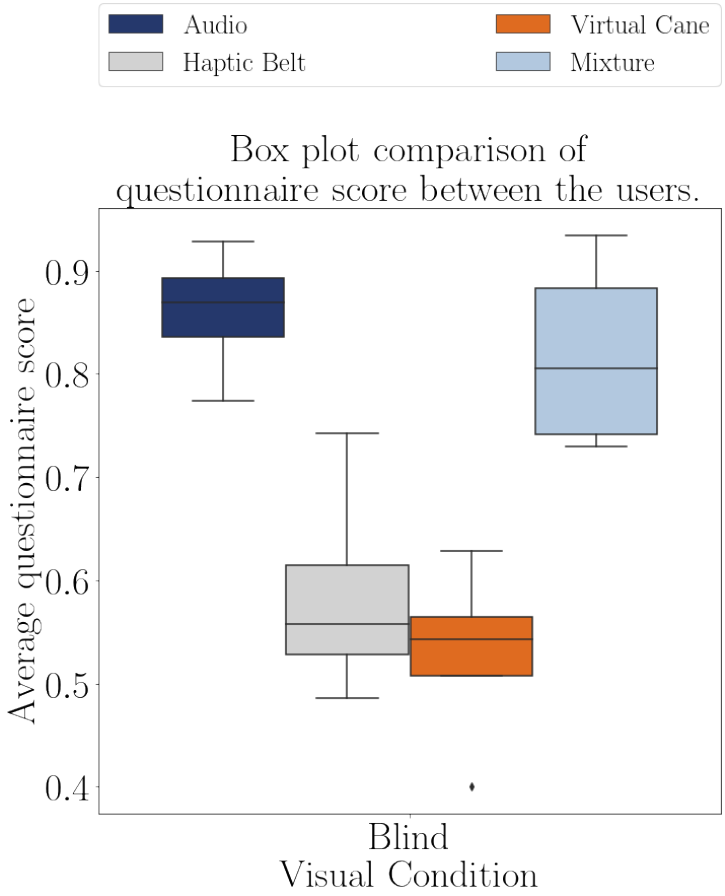

FIGURE 1.23 – Boxplot of the questionaire score of the blind participants grouped by method.

TABLE 1.19 – Anova p-value for the questionnaire score on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (blocks)	0.042	3	0.110	2.014	
Method	0.329	3	0.014	15.677	0.001**
Experimental error	0.063	9	0.007		
Total	0.434	15			

The Table 1.20 presents the conclusion of a pairwise Fisher LSD test of the blind NASA-TLX score between all the guidance methods. The results show that only the "Audio" and "Mixture" have the same statistically result and that there is a difference between the both "Haptic Belt" and "Virtual Cane".

The LSD Table 1.20 confirms the information of the Figure 1.23 that 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 participants 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.

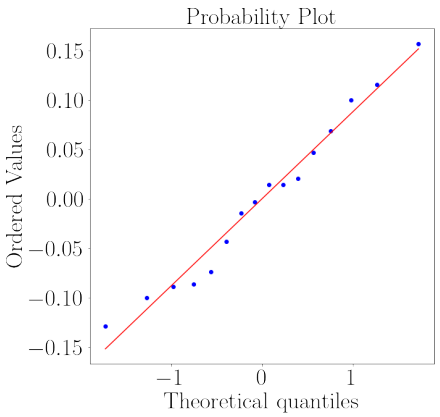

FIGURE 1.24 – QQ plot of the questionnaire score of the blind participants on each method.

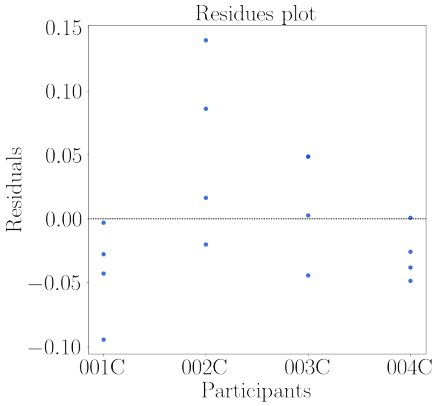

FIGURE 1.25 – Residual plot of the questionnaire score the blind participants on each method.

TABLE 1.20 – Cross validation p-value for the questionnaire score on each method for blinded users.

Method			Analysis
Audio	X	Haptic Belt	$H_1 : \mu_{Audio} \neq \mu_{HapticBelt} **$
Audio	X	Virtual Cane	$H_1 : \mu_{Audio} \neq \mu_{VirtualCane} **$
Audio	X	Mixture	$H_0 : \mu_{Audio} = \mu_{Mixture}$
Haptic Belt	X	Virtual Cane	$H_1 : \mu_{HapticBelt} \neq \mu_{VirtualCane} **$
Haptic Belt	X	Mixture	$H_1 : \mu_{HapticBelt} \neq \mu_{Mixture} **$
Virtual Cane	X	Mixture	$H_1 : \mu_{VirtualCane} \neq \mu_{Mixture} **$

1.1.2 Physiological data

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 slight decrease from the "First" to the "Return" round. The heartrate variance is expected to slight increase from the "First" to the "Return" 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.

1.1.2.1 Electrocardiogram (ECG) data

The ECG analysis is divided into two different types

- Heart rate;

  This analysis checks the heartbeat frequency;

- Heart rate variance.

  This analysis checks the heartbeat frequency variance and it is done by analyzing the variation of the interval between beats.

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 (**??**). 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.

1.1.2.1.1 Analysis of the heartbeat frequency (BPM)

The Table 1.21 presents the average heart rate by each blind participant on each scenes. It is possible to see that the previous expectation cannot be proven, since there is no sistematic pattern in the heartrate variation between the rounds.

TABLE 1.21 – Average BPM felled by the blinded participants.

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

In the Figure 1.26 is plotted the average data presentend in the previoes table. There is a slight increase in the heartrate between the rounds, with the exception of the "Base" method. That means that, in the average, the participants felt more demandful in the "Return" round.

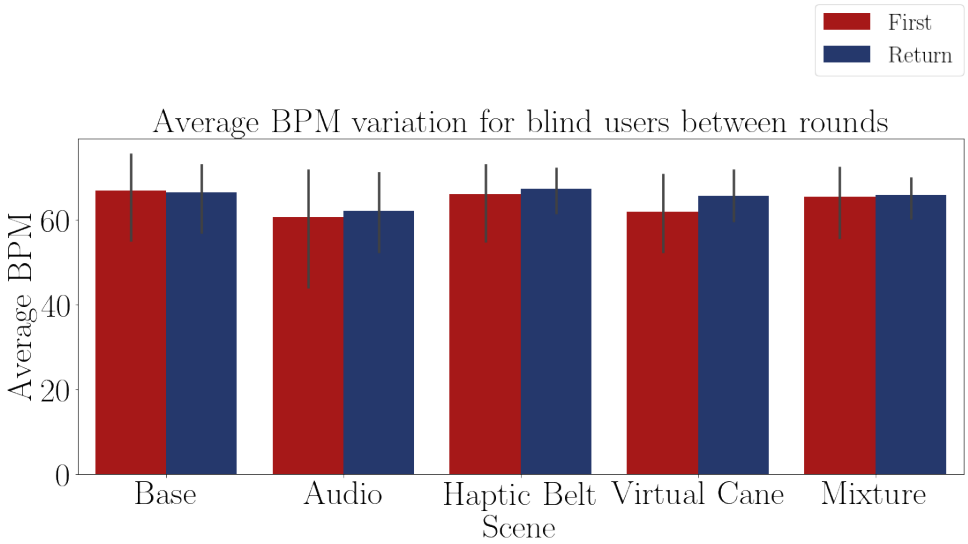

FIGURE 1.26 – Barplot of the average BPM of the blind participants on each method.

The Table 1.22 show the average heartbeat frequency variation between the rounds of each group. As it was shown in the Figure 1.26, only the "Base" method has a negative average variaton between the rounds. It is also posible to see that the Virtual Cane variation was the highest, hence it was also the highest mental workload.

TABLE 1.22 – ECG average BPM for each method of the blind participants.

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	-0.58	1.40	1.09	3.79	0.57

The Figure 1.27 show a comparison between the methods. There is no big difference between them, but it is posible to separate them in two groups based on their similarity. One with "Base", "Haptic Belt" and "Mixture" methods and the other with "Audio" and "Virtual Cane". The Figure 1.28 presents the average heartreate frequency grouped by round.

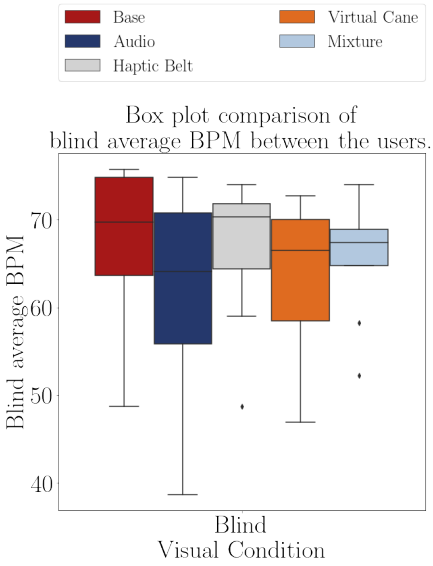

FIGURE 1.27 – Boxplot of the BPM of the blind participants grouped by method.

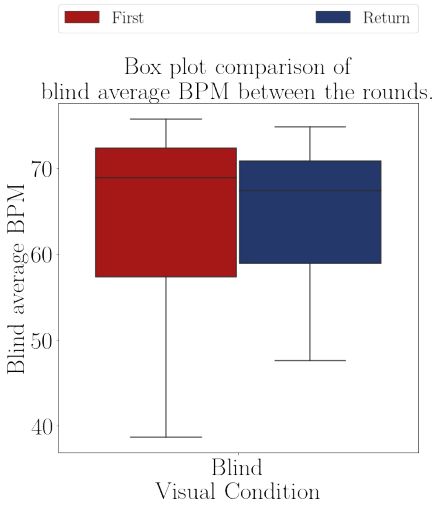

FIGURE 1.28 – Boxplot of the BPM of the blind participants grouped by round.

The Figures 1.29 and 1.30 shows the distribution and variance of the Table 1.21. These Figures shows that the data are normally distributed but the participants had different that the methods have a similar variance. The Table 1.23 shows the ANOVA test p-value of the heart rate frequency of the "blind" sample. The p-value indicates that there is no effect of the methods, rounds and neither their interaction in the heartrate frequency.

According to the ANOVA test at Table 1.23, there is no effect from the method, the round or the interaction between them in the heartrate frequency. It is posible to notice some small difference in the Figure 1.27 but maybe because of the small sample size, it was no sensitive enough to be proved by the ANOVA test. But inside that Figure

TABLE 1.23 – Anova p-value for the BPM on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (Blocks)	2807.274	3	935.758	49.361	
Methods	164.045	4	41.011	2.163	0.100
Rounds	15.693	1	15.693	0.828	0.371
Interaction	20.606	4	5.152	0.272	0.894
Experimental Error	511.853	27	18.958		
Total	3519.471	39			

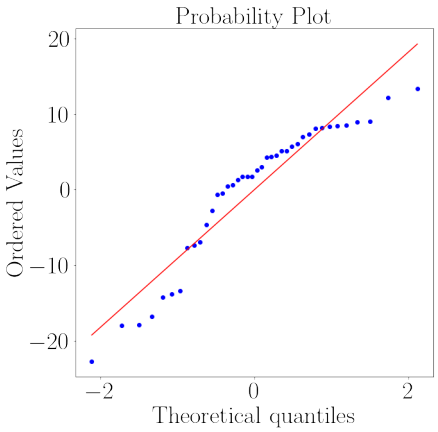

FIGURE 1.29 – QQ plot of the BPM of the blind participants on each method.

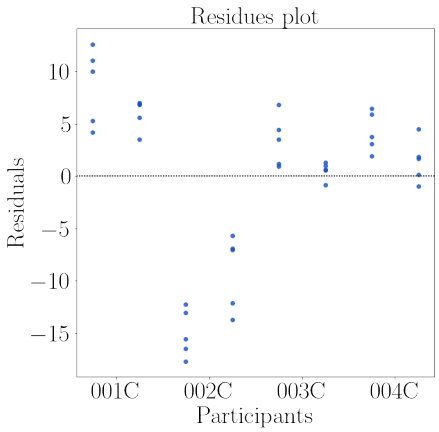

FIGURE 1.30 – Residual plot of the BPM score the blind participants on each method.

1.1.2.1.2 Analysis of the heartbeat variance (SDNN)

The Table 1.24 presents the standard deviation of the interbeat interval by each participant on each scenes. As it was with the Table 1.21, it is not posible to draw a pattern inside this Table. Different participant had increase, or decrease, with different methods.

TABLE 1.24 – Average SDNN of the blind participants during the each round and method.

Participant	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001C	First	81.292	107.061	124.737	163.968	129.054
	Return	120.719	130.885	131.590	157.589	124.786
002C	First	73.761	98.863	81.140	33.977	79.289
	Return	108.940	49.627	42.815	114.057	107.545
003C	First	36.870	38.325	35.101	42.392	43.692
	Return	52.750	41.196	44.256	42.602	46.145
004C	First	70.728	86.827	62.560	85.900	70.472
	Return	71.950	74.895	70.017	66.089	104.040

Inside the barplot Figure 1.31 shows the average SDNN in each method. It is posible to notice that some method had an increase and some a decrease in the SDNN. The ones

that indicate an increase would mean that the participant felt a lesser mental workload in the "Return" round, whilst the deacrese means the opposite.

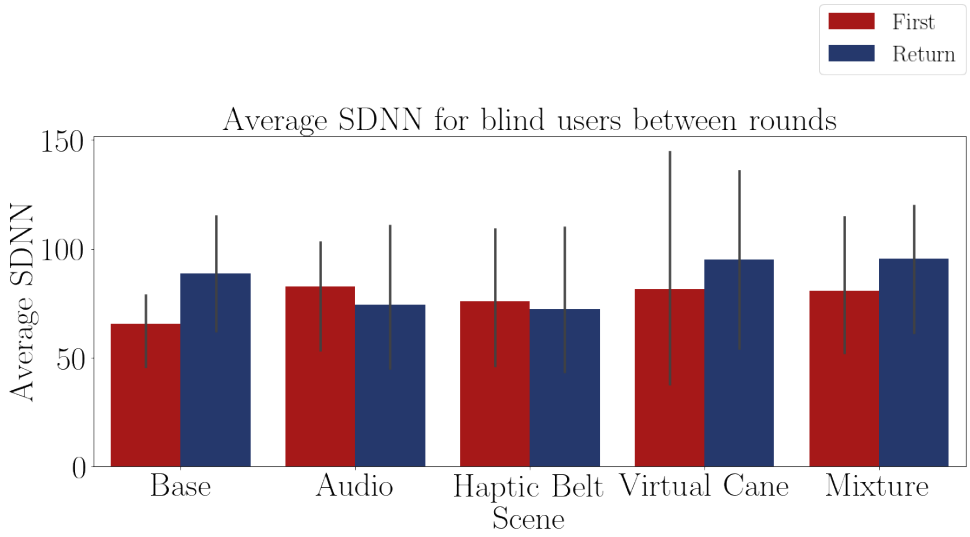

FIGURE 1.31 – Barplot of the average SDNN of the blind participants on each method.

The Table 1.25 presents the average SDNN variation between the rounds. It shows that only the "Audio" and the "Haptic Belt" methods shown a increase in the mental workload.

TABLE 1.25 – ECG average SDNN for each method of the blind participants.

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	22.93	-8.62	-3.72	13.52	15.00

The Figures 1.32 presents the distribution of each method SDNN. It noticeable that the "Base" method has a different SDNN than the rest. The "Virtual Cane" also has a different distribution from the rest. The Figure 1.33 presents the SDNN grouped by the rounds. It shows a slight difference between the rounds.

The Figures 1.34 and 1.35 shows the distribution and variance of the Table 1.24. These Figures shows that the data are normally distributed but the participants had different that the methods have a similar variance. The Table 1.26 shows the ANOVA test p-value of the heartbeat interval variance of the "blind" sample. The p-value indicates that there is no effect of any factor.

The Table 1.26 does not prove that any method or round has some influence in the heartbeat interval variance, thus in the Mental Workload. Although, in the Figure 1.32 it is posible to notice that the "Base" method has a different distribution. As it has already commented before, maybe the result of the anova test is a conseguence of a small sample size.

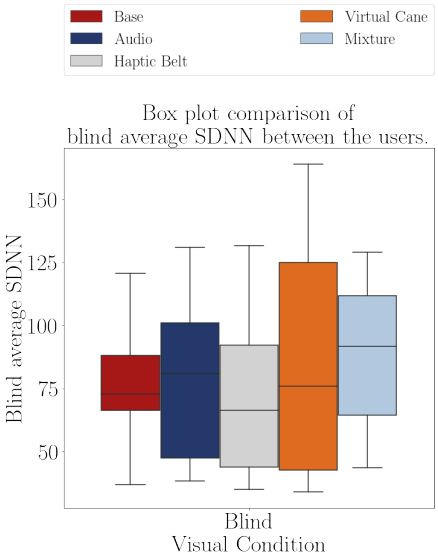

FIGURE 1.32 – Boxplot of the SDNN of the blind participants grouped by method.

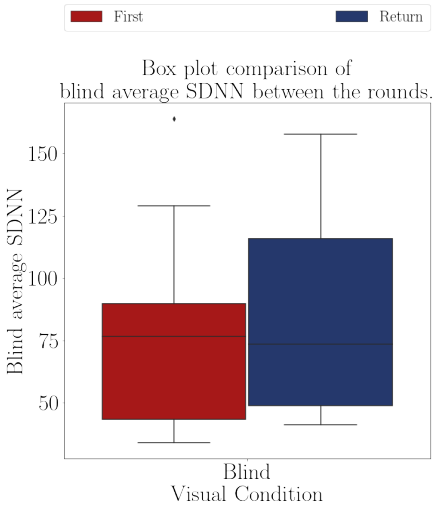

FIGURE 1.33 – Boxplot of the SDNN of the blind participants grouped by round.

TABLE 1.26 – Anova p-value for the average SDNN on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (Blocks)	36520.955	3	12173.652	30.932	
Methods	1394.166	4	348.542	0.886	0.486
Rounds	612.182	1	612.182	1.555	0.223
Interaction	1431.284	4	357.821	0.909	0.473
Experimental Error	10626.244	27	393.565		
Total	50584.831	39			

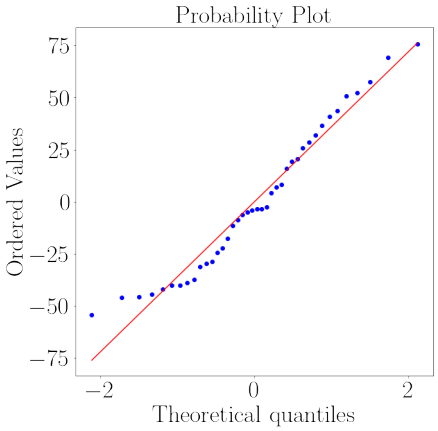

FIGURE 1.34 – QQ plot of the SDNN of the blind participants on each method.

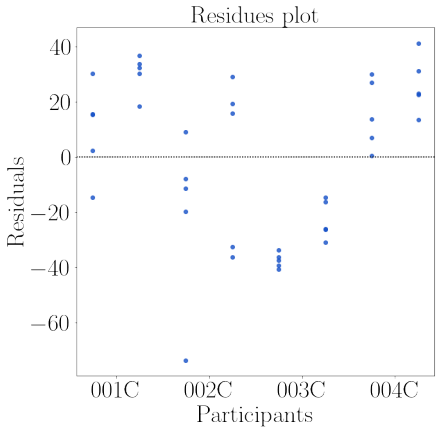

FIGURE 1.35 – Residual plot of the SDNN of the blind participants on each method.

1.1.2.2 Galvanic skin reaction and temperature data;

The GSR analysis is made by analyzing its average variation between the rounds. The bigger the variation, the bigger the stress, arousal or mental workload felt by the

participant. 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.

TABLE 1.27 – Average GSR felled by the participants [μS].

Part.	Visual Condition	Round	Baseline	Base	Audio	Haptic Belt	Virtual Cane	Mixture
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

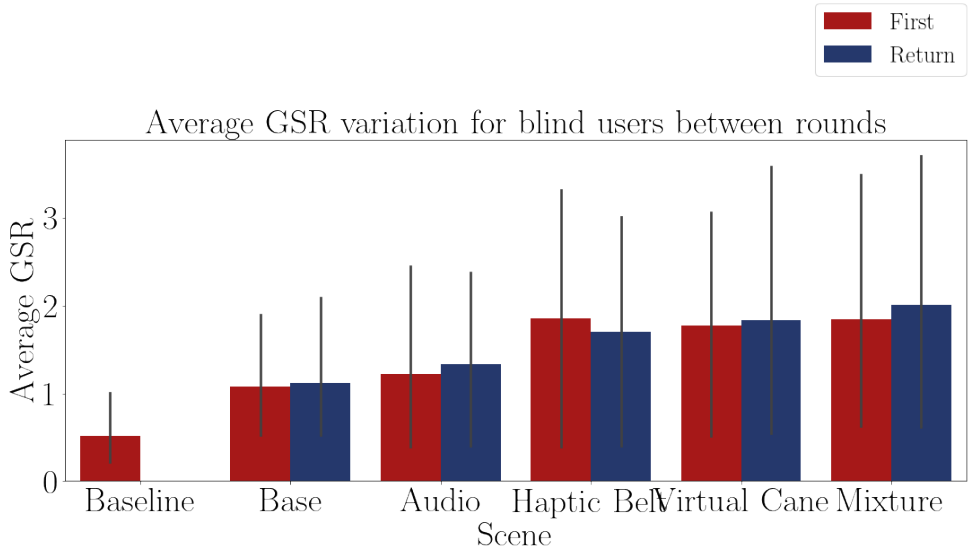

FIGURE 1.36 – Barplot of the average SDNN of the blind participants on each method.

TABLE 1.28 – Average GSR variation by the blind participants

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	141.10	127.32	244.62	307.18	344.366

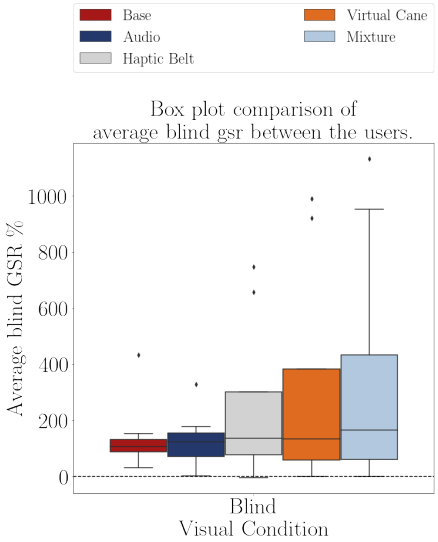

FIGURE 1.37 – Boxplot of the SDNN of the blind participants grouped by method.

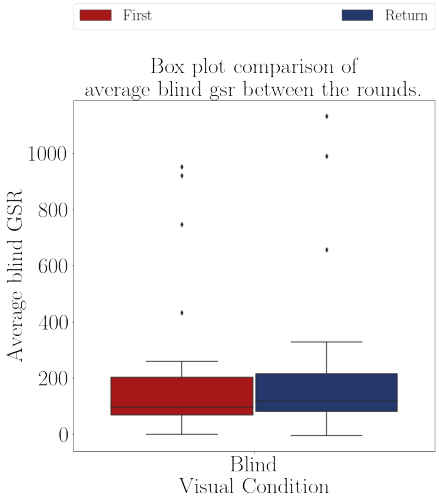

FIGURE 1.38 – Boxplot of the SDNN of the blind participants grouped by round.

TABLE 1.29 – Anova p-value for the GSR score on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Participants (blocks)	1918983.831	3	221654.018	9.552	
Method	886616.073	4	639661.277	3.310	0.048**
Experimental error	803557.209	12	66963.101		
Total	3609157.113	39			

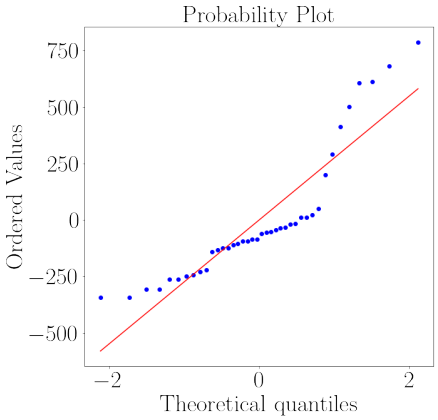

FIGURE 1.39 – QQ plot of the SDNN of the blind participants on each method.

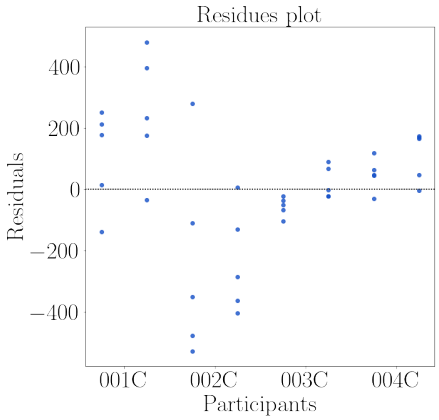

FIGURE 1.40 – Residual plot of the SDNN of the blind participants on each method.

TABLE 1.30 – Cross validation p-value for the GSR on each method for blinded users.

Method			Analysis
Base	X	Audio	$H_0 : \mu_{Base} = \mu_{Audio}$
Base	X	Haptic Belt	$H_0 : \mu_{Base} = \mu_{HapticBelt}$
Base	X	Virtual Cane	$H_1 : \mu_{Base} \neq \mu_{VirtualCane} **$
Base	X	Mixture	$H_1 : \mu_{Base} \neq \mu_{Mixture} **$
Audio	X	Haptic Belt	$H_1 : \mu_{Audio} \neq \mu_{HapticBelt} **$
Audio	X	Virtual Cane	$H_1 : \mu_{Audio} \neq \mu_{VirtualCane} **$
Audio	X	Mixture	$H_1 : \mu_{Audio} \neq \mu_{Mixture} **$
Haptic Belt	X	Virtual Cane	$H_0 : \mu_{HapticBelt} = \mu_{VirtualCane}$
Haptic Belt	X	Mixture	$H_0 : \mu_{HapticBelt} = \mu_{Mixture}$
Virtual Cane	X	Mixture	$H_0 : \mu_{VirtualCane} = \mu_{Mixture}$

1.2 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.

2 Conclusion

In this final chapter, the goals will be revised along with the results collected. It will be divided into four sections, one for each goal and a final one for future works and suggestions, and each section will have four more subsections, one for each data source gathered and one for a conclusion and commentaries for that goal.

- **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?**

  Based on the gathered data, there was a variation in the mental workload and in the situation awareness during the experiment. This variation show that the users were impacted by the experiment in the virtual reality, but since no experiment outside the virtual reality was made, it is not possible to compare this data and verify that they are similar to one provided by a real scenario.

  Although there was variations inside the experiment, there was also some unexpected results. The heartbeat and the interbeat interval standard deviation did not show the same results as the NASA-TLX indicated. That could be by the fact that the parcipants walked the majority of the time, and that polluted the sensor data, leaving to unrelatable results. It could also be caused because the experiment was made using a virtual reality, and this may have "relaxed" the participants.

  As for the limitations, the participants complained about the sound. The integrate headphone of the VIVE HMD did not provide sounds with a quality good enought for they to locate. A common commentarie was "I feel like the sound origin is inside my head", which was not true. But this can be solved by placing a real sound source in the real environment and use the HMD only for geolocalizing the participant inside de virtual environment.

  Another limitation is the real time position of the furniture. More than once, after a "First round" a furniture was not well aligned with the its virtual model. That caused some frustation on the participant as well in the researchers that had to stop the experiment to fix their position. A solution for this it would be to install real time locator on each piece of furniture.

- **Do non-BVI users, when deprived from their vision, evaluate assistive devices in a similar way as BVI users?**

  - Answers based on the simulation data

    Results from the simulation data and the T-Test showed that the only time data that was different between the groups is the "Audio". Analyzing the rest of the data one can conclude that the results had no difference.

    Graphically it is possible to notice a rather similar average duration between the two groups going along with the conclusion from the T-Test, but there is the matter of the unreliability of this data mentioned before.

  - Answers based on the subjective data

    The T-Test of each questionnaire showed that there are no differences between the groups, but the graphically is noticeable a difference between the groups. These unmatched results may be because of the small sample number.

  - Answers based on the physiological data

    The Figures indicate that there are the groups may have a similar average, but the variation of them are different in most cases. All the T-Tests indicate that both groups have the same variation of workload and arouse.

  - Final conclusions and comentaries

    The T-Test results showed in general that both groups had similar results, while some graphics showed the opposite. This happened maybe because of two reasons. First because of a small sample size. Second because of a tendency of the "sight" sample. The sighted participant all were used to technology and volunteering for experiments, while the same can not be said for the BVI participants.

2.1 Future works and suggestions

For future works related to this one it could be suggested:

- Repeat the experiment in a real situation and compare it with this one to verify the first goal;

  This experiment was made exclusivily using virtual reality, hence it is not posibly to verify the efficiency or the quality of a experiment made using Virtual Reality.

- Repeat the experiment not using the sound from the HMD;

  The BVI users complained about the VIVE HMD sound system. They got confused sometimes and could figure it out if a sound source was coming from forwards or

backwards. One alternative for this problem is to add a physical sound source at each point in the real environment where it was supose to be in the virtual environment. It still related to the virtual reality but it is more realistic.

- Repeat the experiment with bigger sample size and a more diverse sample to verify if the results of the hypothesis test do remain the same;

  As commented before, most ANOVA tests showed one result and the figures showed a different conclusion. This happend because of the small sample size. If the sample size was bigger, maybe both conclusions would be the same.