1 Results' analysis and discussion

Throughout the experiment, three data sources were gathered from the participants, and this chapter will show their values, will explain the process to analyze the data and will discuss their results. Each source will have its section, making up to three sessions, and they are:

- Data collected from the simulation;
- Data collected from questionnaires;
- Data collected from physiological sensors.

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

The processing of each data collected is rather similar and follows these steps:

1. Separate the Blind sample and the Sight sample;
2. Check if the samples are normally distributed;
   If the data is normally distributed then it is possible to use other statistical analyses and verify the results statistically.
3. Check if the "blind" sample is statistically different then the "sight" sample;
   This is one of the goals. To verify that the workload and the situation awareness of the blind participants are different from the sighted participants
4. Calculate the average of of each participant in each method;
5. Calculate the average of the participant group in each method.

1.1 Data from the simulation

Unity3D was programmed to record the time that each user spent in each scene. It is expected that the time analysis will show the following observation:

- The scene made with the white cane would be the fastest and with the less number of impacts;
  Since the participant is already used to this method, it is safe to assume that with the others methods the participant would go slower and hit more furniture on the way.

- Comparing both scenes made with the same method, the second one would have the fastest and with less impact;
  Not only this is expected but also is the intention on having two scenes with each method.

1.1.1 Time elapsed on each scene

The data collected from the participants are shown in the Table 1.1.

TABLE 1.1 – Duration grouped by participant and guidance method (in minutes).

Participant	Visual Condition	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	10:18	13:05	6:42	6:52	7:54
		Return	12:38	6:25	7:41	10:28	5:21
001C	Blind	First	2:11	6:00	10:41	9:02	7:42
		Return	11:21	7:41	6:06	5:36	6:10
002C	Blind	First	2:02	6:17	4:32	7:34	4:08
		Return	13:32	8:06	8:02	3:35	3:57
003	Sight	First	8:06	2:14	2:51	4:21	8:11
		Return	4:11	15:25	6:50	5:25	4:18
003C	Blind	First	2:40	11:16	8:04	5:20	5:42
		Return	6:38	4:59	4:00	8:52	5:32
004	Sight	First	2:40	11:16	8:04	5:20	5:42
		Return	6:38	4:59	4:00	8:52	5:32
004C	Blind	First	2:30	6:26	4:23	5:04	3:54
		Return	8:29	6:14	11:25	4:29	6:24
005	Sight	First	2:33	6:58	5:34	5:09	7:52
		Return	8:16	8:46	4:25	6:45	3:00

The Table 1.2 show the the average time of each participant on each method and they are plotted in the Figure 1.1 and 1.2. The Figure shows that there is no pattern in the relationship between the difference in the rounds of each method with the visual condition of the users.

The Table 1.7 show the the average time of each participant and Figure 1.3 these data

TABLE 1.2 – Average duration grouped by participant and guidance method (in minutes).

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	11:28	9:45	7:12	8:40	6:38	Sight
001C	6:46	6:50	8:23	7:19	6:56	Blind
002C	7:47	7:11	6:17	5:34	4:02	Blind
003	6:09	8:50	4:51	4:53	6:14	Sight
003C	4:39	8:08	6:02	7:06	5:37	Blind
004	4:39	8:08	6:02	7:06	5:37	Sight
004C	5:30	6:20	7:54	4:46	5:09	Blind
005	5:25	7:52	4:59	5:57	5:26	Sight

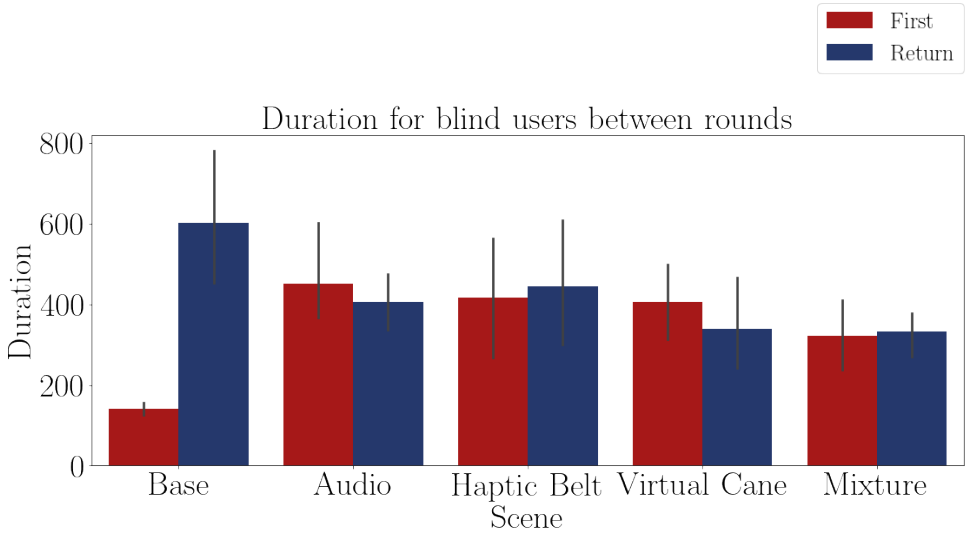

FIGURE 1.1 – Bar plot of the average time of the blind participants on each method.

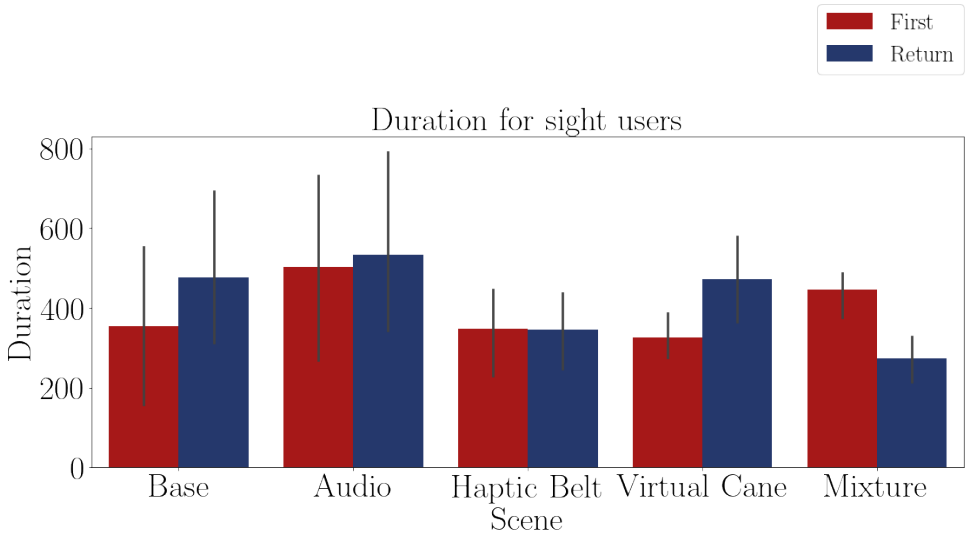

FIGURE 1.2 – Bar plot of the average time of sighted participants on each method.

is plotted. The Figure shows that there could be some difference in the time between the methods, but that would only be assured with a hypothesis test.

TABLE 1.3 – Duration difference grouped by participant and guidance method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	22.6%	-50.9%	14.7%	52.4%	-32.2%	Sight
001C	419.6%	28.2%	-42.9%	-37.9%	-20.0%	Blind
002C	563.4%	28.9%	77.3%	-52.6%	-4.2%	Blind
003	-48.3%	587.1%	139.6%	24.5%	-47.5%	Sight
003C	148.9%	-55.8%	-50.3%	66.3%	-2.8%	Blind
004	148.9%	-55.8%	-50.3%	66.3%	-2.8%	Sight
004C	237.4%	-3.2%	160.5%	-11.5%	63.7%	Blind
005	222.8%	25.9%	-20.6%	30.9%	-61.9%	Sight

The Table 1.4 show the the average time grouped by visual condition and Figure 1.4 these data is plotted. The table shows a noticeable difference between the two groups. The Figure 1.4 shows that the global average of the groups in all scenes were almost the same.

TABLE 1.4 – Duration difference grouped by participant and visual Condition.

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	342.3%	-0.5%	36.1%	-8.9%	9.2%
Sight	86.5%	126.6%	20.8%	43.5%	-36.1%

For more correct analysis, one should use statistical methods to analyze. So hypothesis tests were used, but the first step in this analysis is to check if the sample has a normal distribution.

The Table 1.5 shows the Shapiro Wilk test p-value. If this value is higher than 0.05, then the sample is normally distributed. The table 1.5 indicates that the p-values of the time averages are normally distributed hence the steps that follow are allowed to be used.

The Table 1.6 shows the T-test p-value between the time average of the blind sample and the time average of the sight sample. If this value is higher than 0.05, it means that there is no statistical differences between the samples and that both samples had the same time performance. The table 1.6 indicates the time of both the blind and the sighted users are statistically the same, with an exception of the "Audio" method.

The Table 1.8 shows the Anova test p-value of the blind time averages between the guidance methods presented in the Table 1.7. If this value is higher than 0.05, there is at least one method that has no statistical difference between one from the other methods. The table 1.8 indicates that there is no similar variations between the methods, which means that all of the time differences noticed in the methods are relevant.

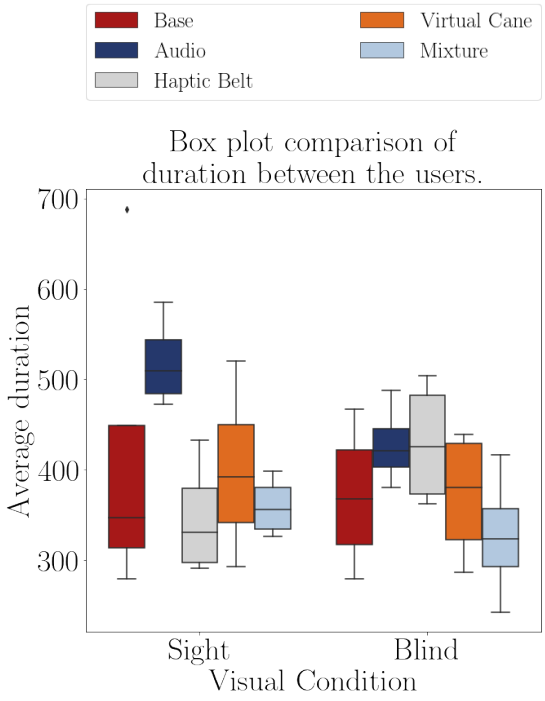

FIGURE 1.3 – Boxplot of the average time of each group on each method.

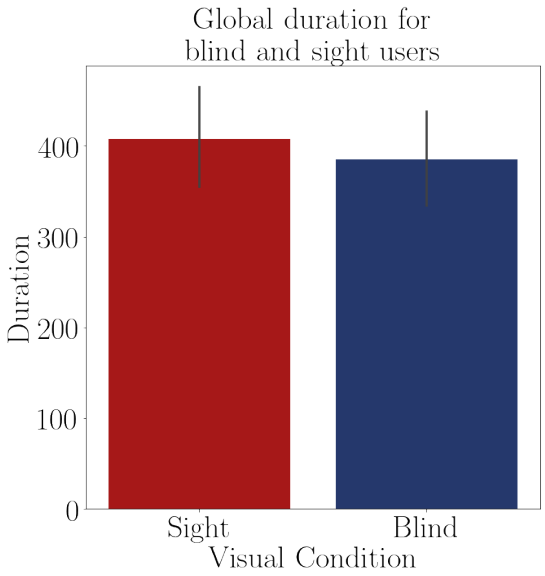

FIGURE 1.4 – Barplot of the average time of each group.

TABLE 1.5 – Shapiro test p-value for the duration of participant in each method.

Method	Shapiro P-Value
Audio blind	0.848
Audio sight	0.623
Haptic Belt blind	0.296
Haptic Belt sight	0.420
Virtual Cane blind	0.402
Virtual Cane sight	0.954
Mixture blind	0.966
Mixture sight	0.619

TABLE 1.6 – T test p-value for the duration for blinded users versus sighted users.

Method	T-Test P-Value
Base	0.675
Audio	0.036
Haptic Belt	0.134
Virtual Cane	0.667
Mixture	0.442

Considering the on Table 1.6, the duration of the "sight" sample is similar to the "blind" sample and considering the conclusion from the ANOVA test and the Figure 1.3, the method that had the better time efficiency was the one mixing all of the methods together, and the least one was the "Audio" method.

Despite all these results above, it is noticeable some outliers in the data, especially in the first participants, when the most minor procedure errors, such as the one to stop the simulation, hence stopping the timer, had happened.

TABLE 1.7 – Duration difference grouped by participant and guidance method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	22.6%	-50.9%	14.7%	52.4%	-32.2%	Sight
001C	419.6%	28.2%	-42.9%	-37.9%	-20.0%	Blind
002C	563.4%	28.9%	77.3%	-52.6%	-4.2%	Blind
003	-48.3%	587.1%	139.6%	24.5%	-47.5%	Sight
003C	148.9%	-55.8%	-50.3%	66.3%	-2.8%	Blind
004	148.9%	-55.8%	-50.3%	66.3%	-2.8%	Sight
004C	237.4%	-3.2%	160.5%	-11.5%	63.7%	Blind
005	222.8%	25.9%	-20.6%	30.9%	-61.9%	Sight

TABLE 1.8 – Anova p-value for the duration of each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Between factors	360193.228	4	90048.307	8.921	0.001
Inside factors	151418.029	15	10094.535		
Total	511611.257	19			

1.2 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. Also that there is a noticeable difference between the sight sample mental workload and the blind sample 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 and a difference between the "blind" sample and the "sight" sample.

- Guidance method's questionnaire.

  Meant to assess the user experience with each method.

1.2.1 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.

1.2.1.1 Analysis of the mental demand scale

The Table 1.9 presents the mental demand averages by each participant on each scenes and they are plotted in the Figures 1.5 and 1.6.

TABLE 1.9 – Mental demand felled by the participants.

Participant	Visual Condition	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	6	12	11	5	9
		Return	6	13	13	5	10
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
003	Sight	First	2	18	18	16	10
		Return	1	12	15	11	8
003C	Blind	First	5	5	5	8	1
		Return	3	1	1	2	1
004	Sight	First	8	17	20	12	20
		Return	5	12	15	10	15
004C	Blind	First	9	10	15	10	10
		Return	7	10	14	8	10
005	Sight	First	2	4	12	10	13
		Return	2	6	10	6	12

The Table 1.10 show the the average mental demand between the rounds of each participant and the Figure 1.7 these data is plotted. The figure shows a noticeable difference between the two groups. The Figure 1.7 indicates a visual difference between the mental demand felt by the sighted participants and the mental demand felt by the blind participants. Inside the blind groups is also noticeable a difference between the methods, but the ones that are different do not show a better performance, instead of higher mental demand than the one felt during the "Base" method.

The Table 1.11 show the the average mental demand grouped by visual condition and these data is plotted in Figure 1.8. Both the table and the figure also show the difference between the mental demand of the "sight" sample and the "blind" sample.

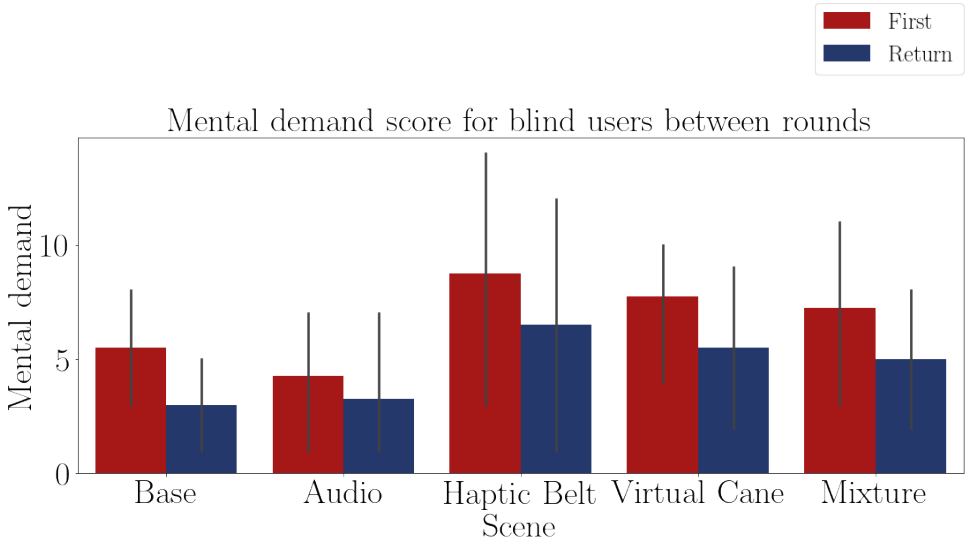

FIGURE 1.5 – Bar plot of the average mental demand of the blind participants on each method.

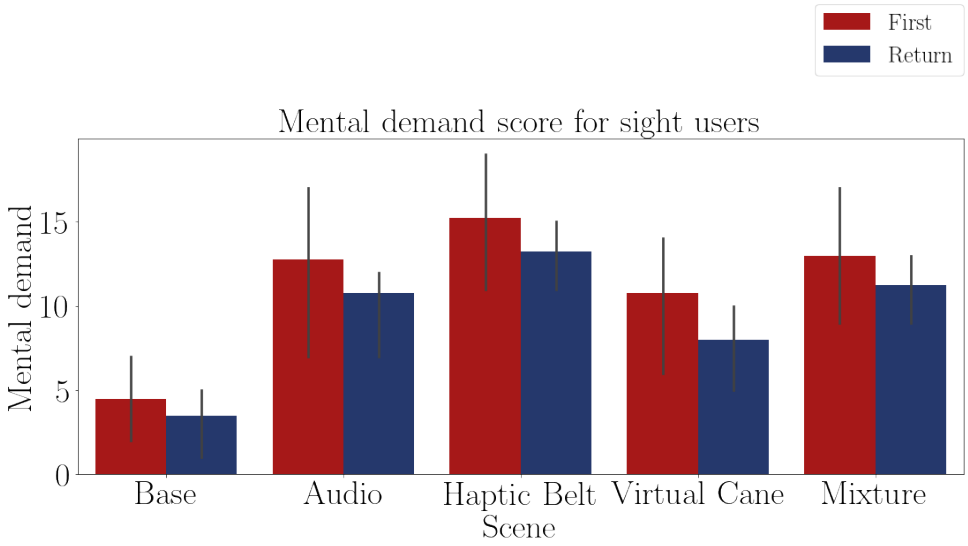

FIGURE 1.6 – Bar plot of the average mental demand of the sighted participants on each method.

TABLE 1.10 – Mental demand average by participant and method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	6.00	12.50	12.00	5.00	9.50	Sight
001C	2.00	1.00	12.00	2.50	6.00	Blind
002C	3.00	1.00	1.00	10.00	7.50	Blind
003	1.50	15.00	16.50	13.50	9.00	Sight
003C	4.00	3.00	3.00	5.00	1.00	Blind
004	6.50	14.50	17.50	11.00	17.50	Sight
004C	8.00	10.00	14.50	9.00	10.00	Blind
005	2.00	5.00	11.00	8.00	12.50	Sight

TABLE 1.11 – 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

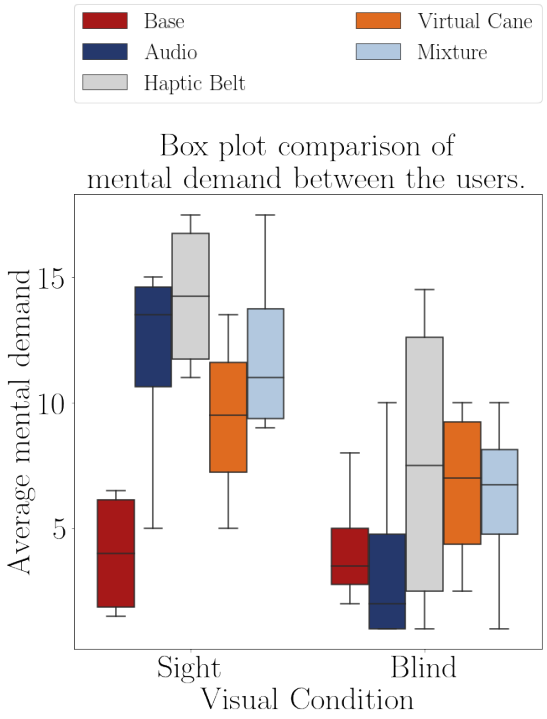

FIGURE 1.7 – Boxplot of the average mental demand of participant.

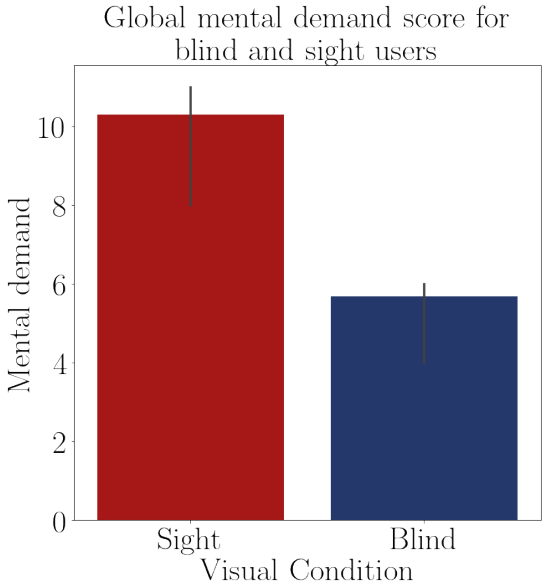

FIGURE 1.8 – Barplot of the average mental demand of each group.

The Shapiro–Wilk normality test on the Table 1.23 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 1.13, the only method that showed a difference in the mental demand between the "sight" sample and the "blind" sample is the audio method. In the other methods both samples had a similar mental demand.

The Table 1.14 shows the Anova test p-value of the mental demand average of the "blind" sample between the guidance methods presented on the Table 1.10. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.

The Table 1.15 presents the conclusion of a pairwise Fisher LSD test of the blind mental demand average between all the guidance methods. The results show that only the "Haptic Belt" caused a different mental demand average than the one noticed on the

TABLE 1.12 – Shapiro test p-value for the mental demand for each method and visual condition.

Method	Shapiro P-Value
Base blind	0.369
Base sight	0.145
Audio blind	0.066
Audio sight	0.117
Haptic Belt blind	0.346
Haptic Belt sight	0.300
Virtual Cane blind	0.555
Virtual Cane sight	0.948
Mixture blind	0.771
Mixture sight	0.339

TABLE 1.13 – T test p-value for the mental demand on each method for blinded users versus sighted users.

Method	T-Test P-Value
Base	0.897
Audio	0.044
Haptic Belt	0.122
Virtual Cane	0.320
Mixture	0.070

TABLE 1.14 – Anova p-value for the average mental demand on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Between factors	42.575	4	10.644	0.556	0.698
Inside factors	287.062	15	19.137		
Total	329.637	19			

"Base" Method.

TABLE 1.15 – Cross validation p-value for the average mental demand 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	$H_0 : \mu_{Base} = \mu_{VirtualCane}$
Base	X	Mixture	$H_0 : \mu_{Base} = \mu_{Mixture}$
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	$H_0 : \mu_{Audio} = \mu_{Mixture}$
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	$H_0 : \mu_{VirtualCane} = \mu_{Mixture}$

The Table 1.17 shows the Anova test p-value of the mental demand variation of the "blind" sample between the guidance methods presented on the Table 1.16. 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 1.18 presents the conclusion of a pairwise Fisher LSD test of the blind mental demand variation between all the guidance methods. The results show that the

TABLE 1.16 – Mental demand variation by participant and method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	0.0%	8.3%	18.2%	0.0%	11.1%	Sight
001C	-66.7%	0.0%	-28.6%	-33.3%	0.0%	Blind
002C	-80.0%	0.0%	0.0%	0.0%	-75.0%	Blind
003	-50.0%	-33.3%	-16.7%	-31.2%	-20.0%	Sight
003C	-40.0%	-80.0%	-80.0%	-75.0%	0.0%	Blind
004	-37.5%	-29.4%	-25.0%	-16.7%	-25.0%	Sight
004C	-22.2%	0.0%	-6.7%	-20.0%	0.0%	Blind
005	0.0%	50.0%	-16.7%	-40.0%	-7.7%	Sight

TABLE 1.17 – 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	0.604	0.666
Inside factors	18007.946	15	1200.530		
Total	20909.752	19			

"Virtual Cane" method has a similar mental demand variation to the "Base" method. All other methods have a different variation. This can be seen at the Table 1.19 and in the Figure 1.9 compiles the mental demand average of the methods observed on all of the participants.

TABLE 1.18 – Cross validation p-value for the mental demand variation on each method for blinded users.

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

According to T-Test on Table 1.13, the mental demand of the "sight" sample is similar to the "blind" sample, excluding only the "Audio" method.

According to both Anova test at Table 1.14 and 1.17, Tables 1.10 and 1.16 and Figure

TABLE 1.19 – 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%	-18.750
Sight	-21.9%	-1.1%	-10.0%	-22.0%	-10.395

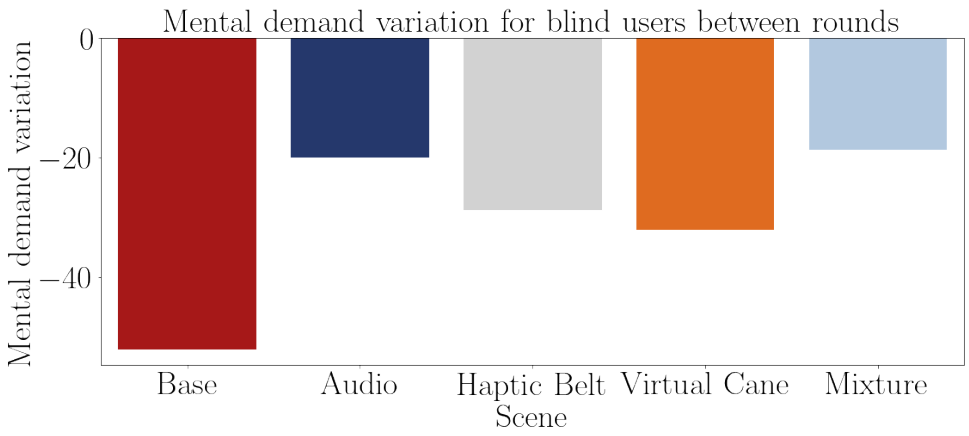

FIGURE 1.9 – Barplot of the average mental demand variation from the blind participants of each method.

1.9, none of the methods did provoke a higher mental demand variation than the one notice on the "Base" method on the participants and the only different mental demand average was noticed on the "Haptic Belt" method, and is higher mental demand than the "Base" method.

1.2.1.2 Analysis of the NASA-TLX score

The Table 1.20 presents the Nasa score averages by each participant on each scenes and they are plotted in the Figures 1.10 and 1.11. It is notible that after each "First" round the Nasa score diminishes for both "sight" and "blind" participants.

The Table 1.21 shows the average Nasa score between the rounds of each participant and in the Figure 1.12 this data is plotted. The table and the figure also show 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.

The Table 1.22 show the the average Nasa score grouped by visual condition and these data is plotted in Figure 1.13. Both the table and the figure also show the difference between the mental demand of the "sight" sample and the "blind" sample.

TABLE 1.20 – NASA score felled by the participants.

Participant	Visual Condition	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
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

TABLE 1.21 – NASA-TLX score grouped by participant and method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	7.9%	10.6%	10.3%	6.6%	9.2%	Sight
001C	4.5%	4.0%	7.8%	4.8%	6.2%	Blind
002C	5.4%	4.8%	4.8%	8.0%	6.1%	Blind
003	4.6%	8.2%	9.9%	8.7%	5.7%	Sight
003C	4.0%	3.9%	4.5%	5.1%	3.5%	Blind
004	6.8%	13.3%	12.8%	11.2%	14.0%	Sight
004C	9.2%	9.6%	12.2%	9.5%	10.9%	Blind
005	5.0%	7.7%	8.8%	7.8%	7.8%	Sight

TABLE 1.22 – 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

The Shapiro–Wilk normality test on the Table **??** shows that these data are normally distributed, with an exception of the "Audio" Nasa score. This means that further analysis

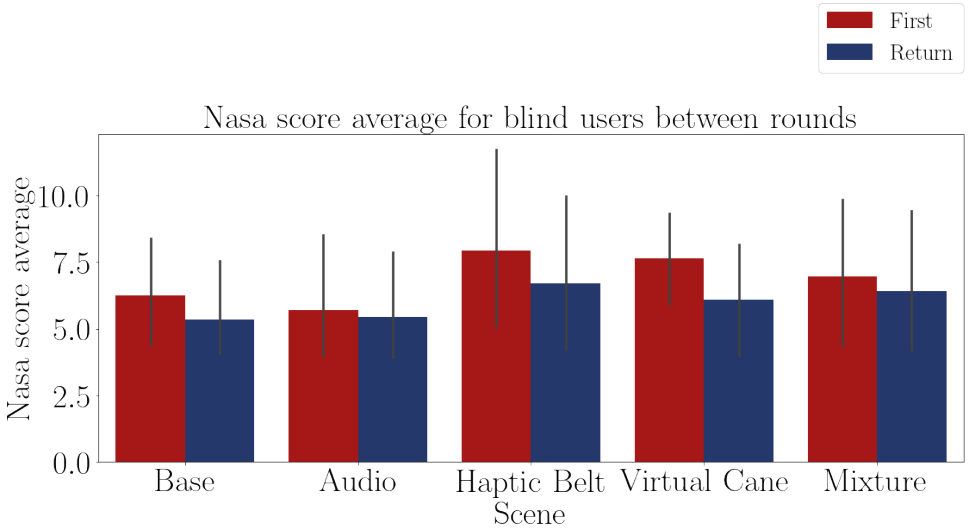

FIGURE 1.10 – Bar plot of the average Nasa-TLX score of the blind participants on each method.

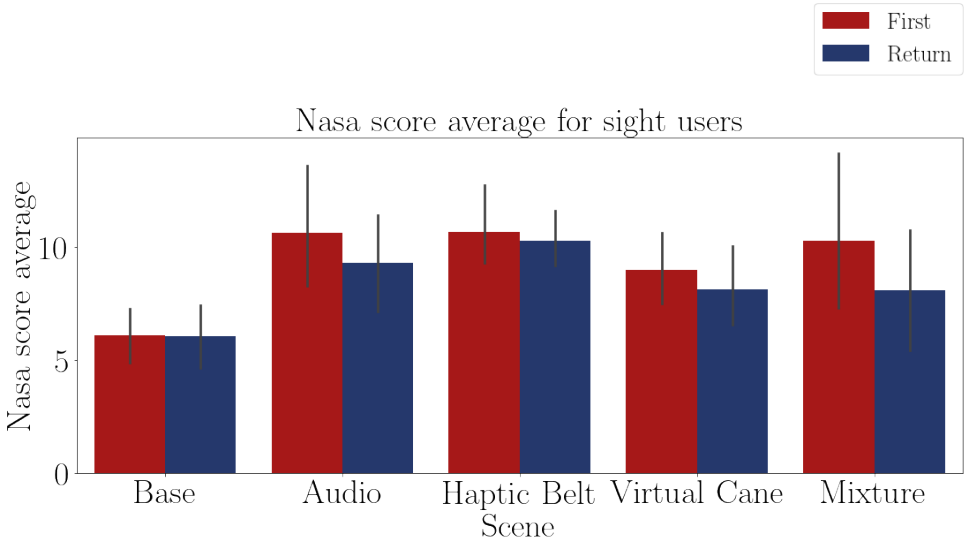

FIGURE 1.11 – Bar plot of the average Nasa-TLX score of the sighted participants on each method.

cannot be applied to this method.

According to the T-Test presented in the Table 1.24 it cannot be verified that the average Nasa score is different between the "sight" and the "blind" samples.

The Table 1.25 shows the Anova test p-value of the average Nasa score, presented in the Table 1.21, 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 1.18 presents the results of a pairwise Fisher LSD test of the blind Nasa score average between all the guidance methods. The results show that all of the averages are statistically the same. That means that there is no difference in the Nasa score felt by the "blind" sample between the methods.

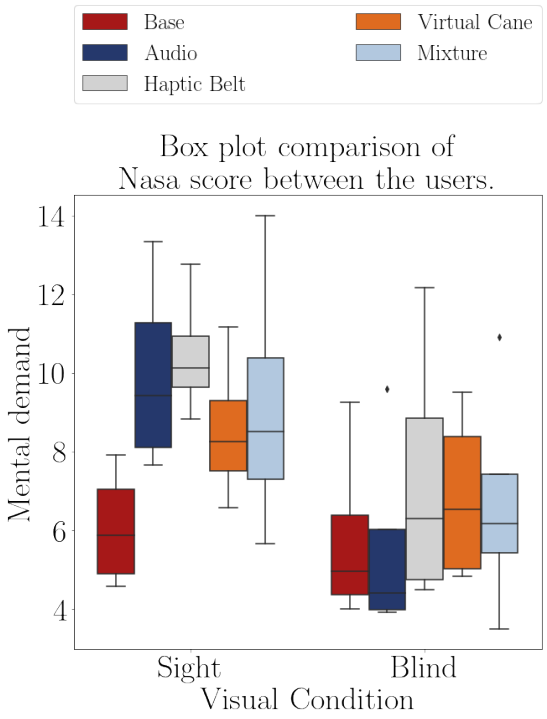

FIGURE 1.12 – Boxplot of the average Nasa-TLX score of the participants.

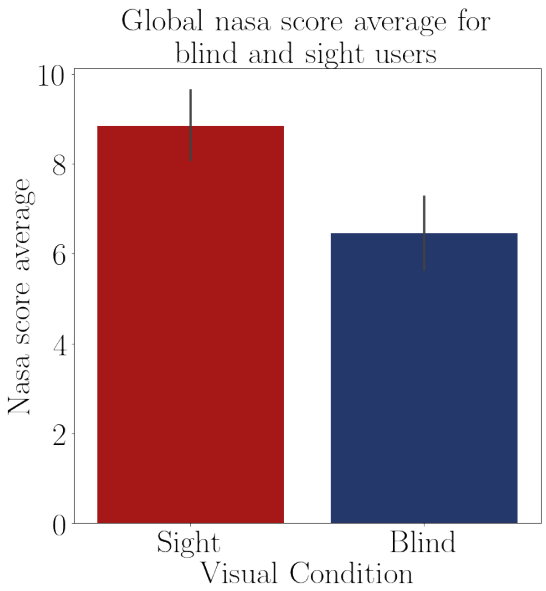

FIGURE 1.13 – Barplot of the average nasa score of each group.

TABLE 1.23 – Shapiro test p-value for the mental demand for each method and visual condition.

Method	Shapiro P-Value
Base blind	0.369
Base sight	0.145
Audio blind	0.066
Audio sight	0.117
Haptic Belt blind	0.346
Haptic Belt sight	0.300
Virtual Cane blind	0.555
Virtual Cane sight	0.948
Mixture blind	0.771
Mixture sight	0.339

TABLE 1.24 – T test p-value for the NASA score on each method for blinded users versus sighted users.

Method	T-Test P-Value
Base	0.855
Audio	0.058
Haptic Belt	0.159
Virtual Cane	0.296
Mixture	0.331

TABLE 1.25 – Anova p-value for the average Nasa score on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Between factors	8.592	4	2.148	0.267	0.895
Inside factors	120.773	15	8.052		
Total	129.365	19			

TABLE 1.26 – Cross validation p-value for the average Nasa score 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	$H_0 : \mu_{Base} = \mu_{VirtualCane}$
Base	X	Mixture	$H_0 : \mu_{Base} = \mu_{Mixture}$
Audio	X	Haptic Belt	$H_0 : \mu_{Audio} = \mu_{HapticBelt}$
Audio	X	Virtual Cane	$H_0 : \mu_{Audio} = \mu_{VirtualCane}$
Audio	X	Mixture	$H_0 : \mu_{Audio} = \mu_{Mixture}$
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	$H_0 : \mu_{VirtualCane} = \mu_{Mixture}$

The Table 1.27 shows the Anova test p-value of the mental demand average of the "blind" sample between the guidance methods presented in the Table 1.30. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.

TABLE 1.27 – Anova p-value for the variation Nasa score 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	1.172	0.362
Inside factors	2639.612	15	175.974		
Total	3464.803	19			

TABLE 1.28 – NASA-TLX score variation grouped by participant and method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	2.1%	8.2%	10.2%	-11.9%	3.7%	Sight
001C	-13.8%	0.0%	-24.5%	-12.9%	-2.6%	Blind
002C	-28.9%	0.0%	0.0%	-22.2%	-26.2%	Blind
003	-10.3%	-32.2%	-4.9%	-17.5%	-25.6%	Sight
003C	0.0%	-4.2%	-31.2%	-47.5%	0.0%	Blind
004	2.5%	-20.2%	-13.4%	-5.8%	-23.2%	Sight
004C	-11.9%	-8.3%	-7.9%	-3.4%	-1.5%	Blind
005	0.0%	0.0%	-3.7%	-4.2%	-37.9%	Sight

The Table 1.29 presents the conclusion of a pairwise Fisher LSD test of the blind Nasa score variation between all the guidance methods. The results show that only the "Audio" method caused a different variation than the one noticed in the "Base" Method. The

Figure 1.14 shows the variation of the Nasa score and one can notice that the variation provoked on the "Audio" method" is a lot lesser than the other ones.

TABLE 1.29 – Cross validation p-value for the variation Nasa score on each method for blinded users.

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

TABLE 1.30 – NASA-TLX score variation grouped by participant and method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	2.1%	8.2%	10.2%	-11.9%	3.7%	Sight
001C	-13.8%	0.0%	-24.5%	-12.9%	-2.6%	Blind
002C	-28.9%	0.0%	0.0%	-22.2%	-26.2%	Blind
003	-10.3%	-32.2%	-4.9%	-17.5%	-25.6%	Sight
003C	0.0%	-4.2%	-31.2%	-47.5%	0.0%	Blind
004	2.5%	-20.2%	-13.4%	-5.8%	-23.2%	Sight
004C	-11.9%	-8.3%	-7.9%	-3.4%	-1.5%	Blind
005	0.0%	0.0%	-3.7%	-4.2%	-37.9%	Sight

According to T-Test on Table 1.24 all of the Nasa scores are similar between both groups.

According to both Anova test at Table 1.25 and 1.25, Table 1.30 1.21 and Figure 1.14, all of the methods have a similar Nasa score average, and the only different variation is provoked by the "Audio" method, and is a smaller variation.

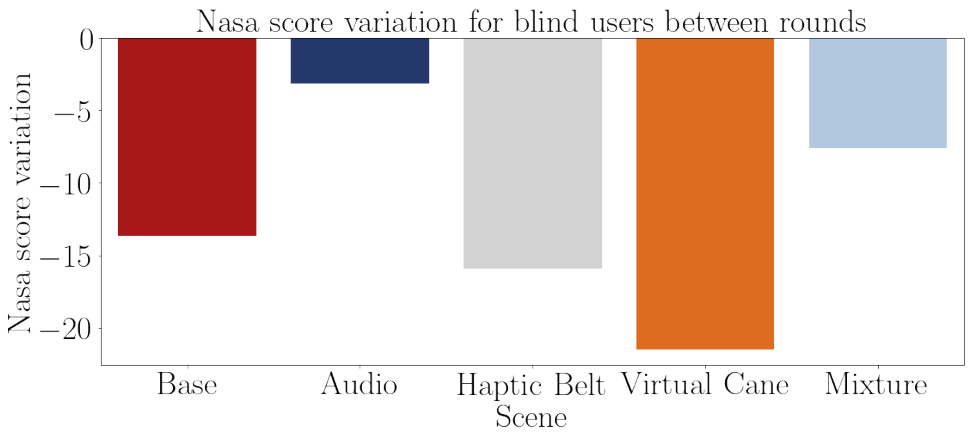

FIGURE 1.14 – Barplot of the Nasa score variation from the blind participants of each method.

1.2.2 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 participant is presented on the Table 1.31 and they are plotted in the Figures 1.15 and 1.16. It is visually noticeable that both of the groups perform better the second time they visit the room.

TABLE 1.31 – Adapted Sagat global score by participant and guidance method.

Participant	Visual Condition	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	1.00	0.45	0.43	0.27	0.650
		Return	1.00	0.60	0.50	0.50	0.450
001C	Blind	First	0.62	0.55	0.53	0.58	0.350
		Return	0.62	0.65	0.85	0.55	0.550
002C	Blind	First	0.68	0.45	0.40	0.45	0.625
		Return	0.53	0.50	0.40	0.65	0.850
003	Sight	First	1.00	0.68	0.60	0.40	0.675
		Return	1.00	0.60	0.72	0.62	0.750
003C	Blind	First	0.72	0.75	0.75	0.47	0.900
		Return	1.00	1.00	0.85	0.90	0.900
004	Sight	First	1.00	0.72	0.80	0.60	0.825
		Return	1.00	0.78	0.95	0.82	0.700
004C	Blind	First	0.75	0.60	0.77	0.50	0.650
		Return	0.90	0.60	0.93	0.72	0.900
005	Sight	First	1.00	0.30	0.32	0.40	0.400
		Return	1.00	0.38	0.30	0.20	0.600

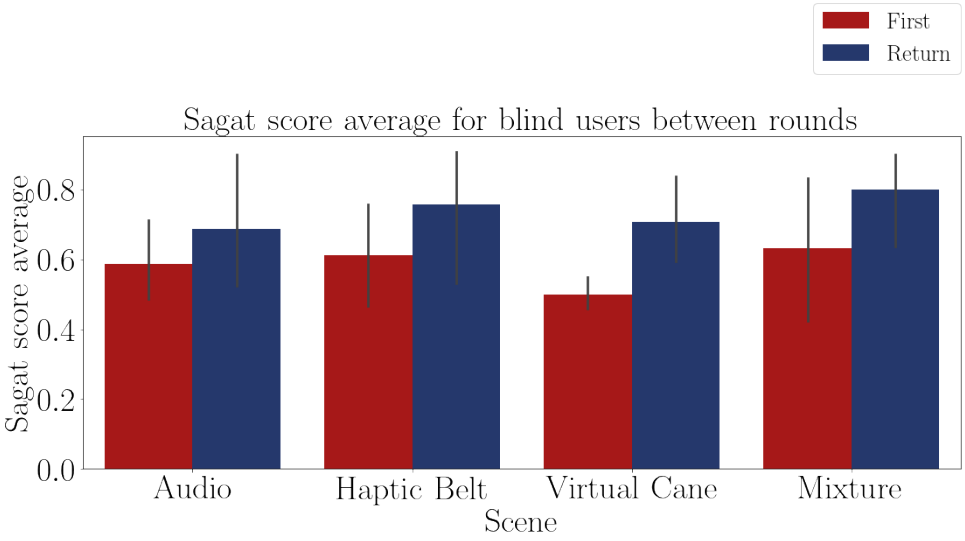

FIGURE 1.15 – Bar plot of the average Sagat score of the blind participants on each method.

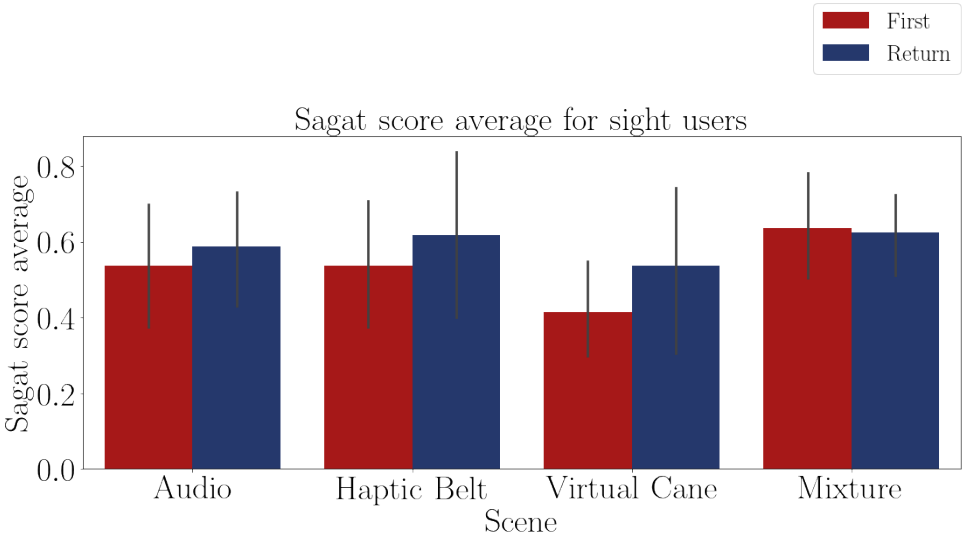

FIGURE 1.16 – Bar plot of the average Sagat score of the sighted participants on each method.

The Table 1.32 shows the average Sagat score between the rounds of each participant and the Figure 1.17 this data is plotted. It is possible only to assume that some methods cause different Sagat scores than others, but both groups performed rather similarly.

The Table 1.33 shows the average Sagat score grouped by visual condition and these data, without considering the "Base" method, is plotted in Figure 1.18. Both the table and the figure also show a slight difference between the score in favor of the "blind" sample.

The Shapiro–Wilk normality test on the Table 1.34 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 1.35, the only method that showed a difference in the Sagat score between the "sight" sample and the "blind" sample is the

TABLE 1.32 – Adapted Sagat average global score grouped by participant and guidance method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	1.00	0.53	0.47	0.38	0.55	Sight
001C	0.62	0.60	0.69	0.57	0.45	Blind
002C	0.60	0.47	0.40	0.55	0.74	Blind
003	1.00	0.64	0.66	0.51	0.71	Sight
003C	0.86	0.88	0.80	0.68	0.90	Blind
004	1.00	0.75	0.87	0.71	0.76	Sight
004C	0.82	0.60	0.85	0.61	0.78	Blind
005	1.00	0.34	0.31	0.30	0.50	Sight

TABLE 1.33 – Adapted Sagat average global score grouped by participant and visual Condition.

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	0.73	0.64	0.68	0.60	0.716
Sight	1.00	0.56	0.58	0.48	0.631

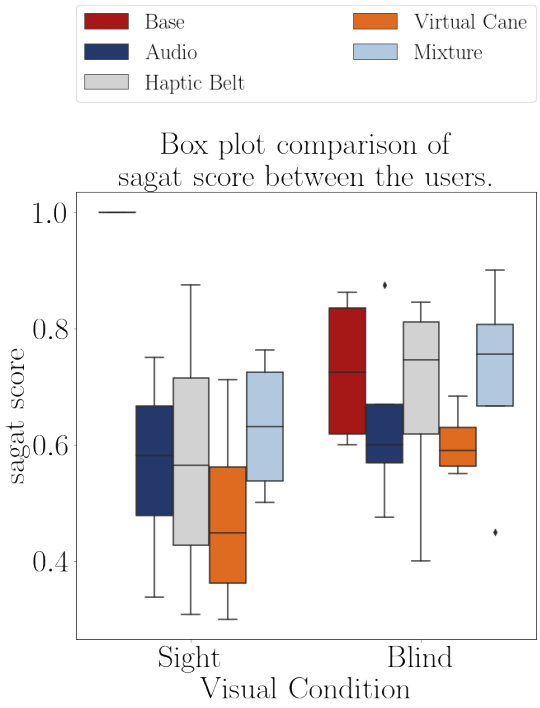

FIGURE 1.17 – Boxplot of the average Sagat score of participant.

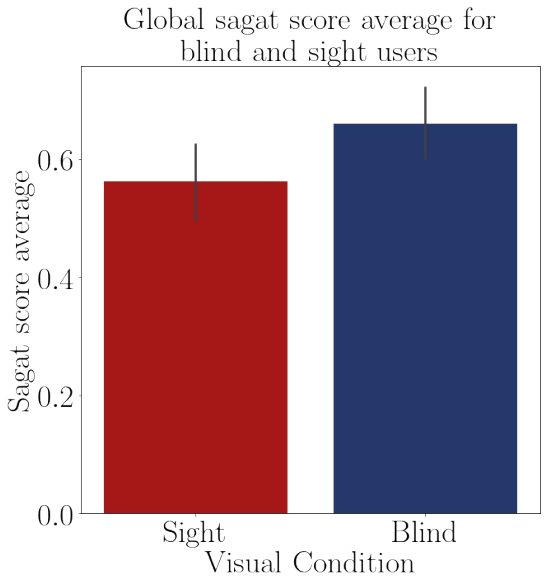

FIGURE 1.18 – Barplot of the average Sagat score of each group.

"Base" method. In the other methods both samples had a similar Sagat score.

The Table 1.36 shows the Anova test p-value of the Sagat score average of the "blind"

TABLE 1.34 – Shapiro test p-value for the Sagat score for each method and visual condition

Method	Shapiro P-Value
Base blind	0.189
Base sight	1.000
Audio blind	0.350
Audio sight	0.925
Haptic Belt blind	0.315
Haptic Belt sight	0.942
Virtual Cane blind	0.549
Virtual Cane sight	0.784
Mixture blind	0.520
Mixture sight	0.446

TABLE 1.35 – T test p-value for the Sagat score on each method for blinded users versus sighted users.

Method	T-Test P-Value
Base	0.007
Audio	0.561
Haptic Belt	0.527
Virtual Cane	0.230
Mixture	0.488

sample between the guidance methods presented in the Table 1.32. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.

TABLE 1.36 – 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	0.045	4	0.011	0.441	0.777
Inside factors	0.380	15	0.025		
Total	0.424	19			

The Table 1.37 presents the analysis of a pairwise Fisher LSD test of the blind average Sagat score between all the guidance methods. The results show that only "Virtual Cane" caused a different average score than the one noticed on the "Base" Method. The rest of the methods did not significantly change it.

TABLE 1.37 – Cross validation p-value for the Sagat score 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	$H_0 : \mu_{Base} = \mu_{Mixture}$
Audio	X	Haptic Belt	$H_0 : \mu_{Audio} = \mu_{HapticBelt}$
Audio	X	Virtual Cane	$H_0 : \mu_{Audio} = \mu_{VirtualCane}$
Audio	X	Mixture	$H_0 : \mu_{Audio} = \mu_{Mixture}$
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	$H_1 : \mu_{VirtualCane} \neq \mu_{Mixture} **$

The Table 1.39 shows the Anova test p-value of the Sagat score variation of the "blind" sample between the guidance methods presented in the Table 1.38. 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 1.38 – Adapted Sagat global score variation grouped by participant and guidance method.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	0.0%	33.3%	15.5%	88.0%	-30.8%	Sight
001C	0.0%	18.2%	59.5%	-5.7%	57.1%	Blind
002C	-22.2%	11.1%	0.3%	44.4%	36.0%	Blind
003	0.0%	-11.1%	21.0%	56.6%	11.1%	Sight
003C	37.9%	33.3%	13.5%	93.1%	0.0%	Blind
004	0.0%	6.9%	18.9%	37.7%	-15.2%	Sight
004C	20.0%	0.0%	20.8%	45.3%	38.5%	Blind
005	0.0%	25.0%	-5.1%	-49.9%	50.0%	Sight

TABLE 1.39 – Anova p-value for the variation Sagat score on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Between factors	3131.542	4	782.885	1.055	0.412
Inside factors	11133.360	15	742.224		
Total	14264.902	19			

The Table 1.40 presents the conclusion of a pairwise Fisher LSD test of the blind Sagat score variation between all the guidance methods. The results show that the "Haptic Belt" and the "Mixture" method have different variations than the "Base" method and they are also different from each other. This can be seen in the Table 1.41 and in the Figure 1.19 compiles the Sagat score average of the methods observed on all of the participants.

According to T-Test on Table 1.35, there is no difference in Sagat score between the "sight" and the "blind" sample.

According to both Anova test at Table 1.36 and LSD test at Table 1.37 only the "Virtual Cane" method has a different Sagat score average and according to the Anova test at Table 1.39 and the LSD test at Table 1.40 the "Virtual Cane" also has a different variation than the "Base" method, with the "Virtual Cane" having a higher, and positive, variation.

Finally, also according with Anova test at Anova test at Table 1.36 and LSD test at Table 1.37 the "Mixture" method also has a significant increase, different the the "Base" method. This increase is also bigger and higher.

TABLE 1.40 – Cross validation p-value for the variation Sagat score 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	$H_0 : \mu_{Audio} = \mu_{HapticBelt}$
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	$H_0 : \mu_{HapticBelt} = \mu_{Mixture}$
Virtual Cane	X	Mixture	$H_0 : \mu_{VirtualCane} = \mu_{Mixture}$

TABLE 1.41 – 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	32.901
Sight	0.00	13.53	12.59	33.12	3.798

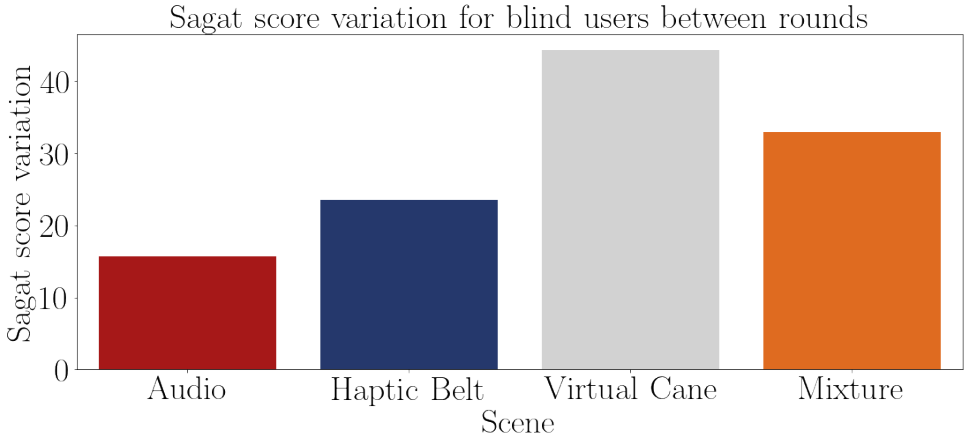

FIGURE 1.19 – Barplot of the average Sagat score variation from the blind participants of each method.

1.2.3 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. Each question was evaluated to favor with a higher score the methods that brought more satisfaction to the user. The Table 1.42 shows the average score of each method and they are plotted in the Figures 1.5 and 1.6

TABLE 1.42 – Guidance method questionnaire average score grouped by participant.

Participant	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	0.46	0.60	0.50	0.56	Sight
001C	0.63	0.71	0.46	0.85	Blind
002C	0.86	0.91	0.49	0.72	Blind
003	0.76	0.71	0.68	0.87	Sight
003C	0.69	0.74	0.54	0.76	Blind
004	0.86	0.77	0.57	0.64	Sight
004C	0.60	0.66	0.40	0.61	Blind
005	0.61	0.74	0.54	0.73	Sight

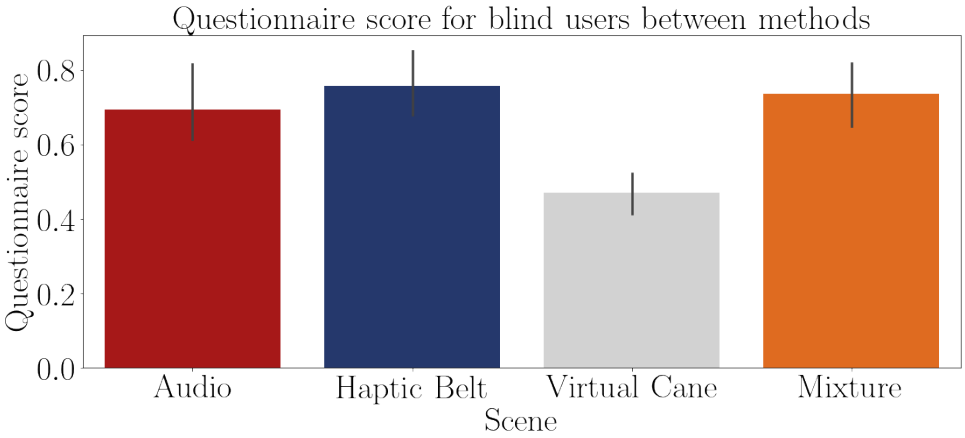

FIGURE 1.20 – Bar plot of the average mental demand of the blind participants on each method.

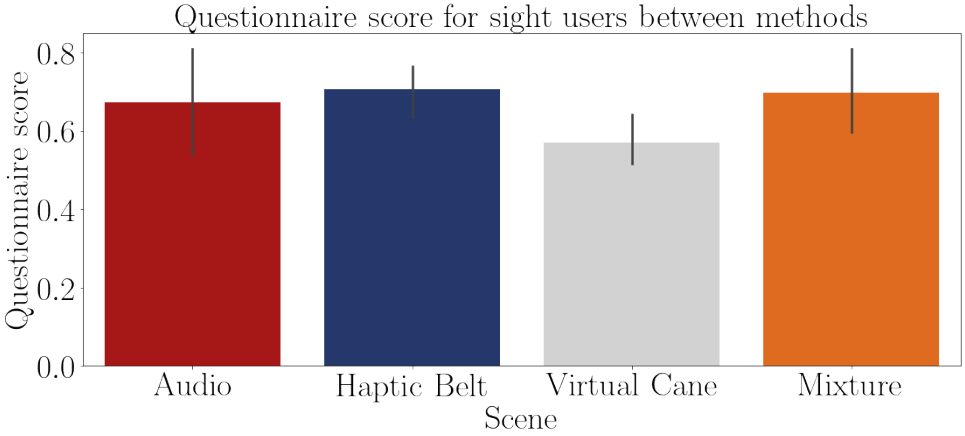

FIGURE 1.21 – Bar plot of the average mental demand of the sighted participants on each method.

The Table 1.43 show the the average questionnaire score of each participant and the Figure 1.22 these data is plotted. It is possible only to assume that some methods cause different Sagat scores than others, but both groups performed rather similarly.

TABLE 1.43 – Guidance method questionnaire average score grouped by visual condition.

Visual Condition	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	0.69	0.76	0.47	0.74
Sight	0.67	0.71	0.57	0.70

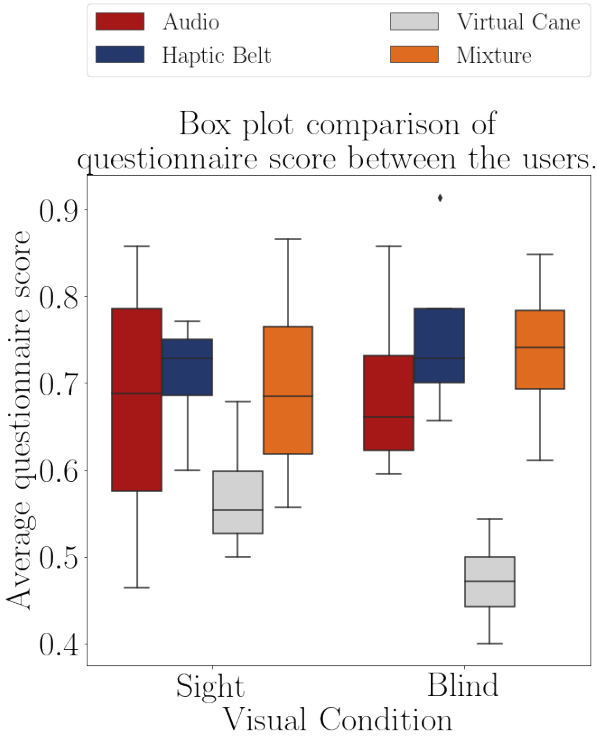

FIGURE 1.22 – Boxplot of the average questionnaire score of each group.

The Shapiro–Wilk normality test on the Table 1.44 shows that these data are normally distributed, with a p-value higher than 0.05, then it is possible to perform the following test to check if there is a significant difference between the methods

The Table 1.45 shows the Anova test p-value of the Sagat score average of the "blind" sample between the guidance methods presented in the Table 1.42. The p-value indicates that all scores are significantly different from each other. That means that the highest scores shown in Table 1.43, which are the "Haptic Belt" and the "Mixture" methods were the most favorite by the participant.

TABLE 1.44 – Shapiro test p-value for the questionnaires score for each method and visual condition.

Method	Shapiro P-Value
Audio blind	0.400
Audio sight	0.882
Haptic Belt blind	0.414
Haptic Belt sight	0.369
Virtual Cane blind	0.995
Virtual Cane sight	0.577
Mixture blind	0.966
Mixture sight	0.925

TABLE 1.45 – 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	0.207	3	0.069	7.080	0.005
Inside factors	0.117	12	0.010		
Total	0.324	15			

1.3 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 only the eliminate a possible increase in the GSR sensor caused by the increase in the temperature. These were all used to assess Mental Workload.

- Electrocardiogram (ECG) data;

  Is expected that the ECG frequency to increase at every "First" round and then a slight decrease in the next round. Also, the variation 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.

1.3.1 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 state induced state by the scene.

After the data gathering, an algorithm in python was used to read the data and separate it accordingly to each participant, method and round. Since the participants moved during the whole experience a lot of noise was collected by the sensors, so these outliers were removed. The following steps were to normalize the data between -1 and 1 and then a peak detection method was used then, if the results were appropriate, the interval between each peak was calculated and saved to be used in the next software. This judgment was made by analyzing the plotted ECG signal and the detected peaks. If the detected peaks are not aligned with the peaks of the signal, then the method's parameters were tuned to fit the detected peaks with the signals' peaks.

The next used software was Kubios HRV Standard. Kubios 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 saved intervals was analyzed and the results were saved in a report file to be read in python again. In python the results were plotted, tabled and statistically tested as the other data. In Appendix **??** 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.

1.3.1.1 Analysis of the heartbeat frequency

The Table 1.46 presents the average heart rate by each participant on each scenes and they are plotted in the Figures 1.23 and 1.24. It is possible to see that 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.

The Table 1.47 show the the average Sagat score between the rounds of each participant and the Figure 1.17 these data is plotted. It is possible only to assume that some methods cause different Sagat scores than others, but both groups performed rather similarly.

The Figure 1.25 shows a comparison between both groups

The Table 1.48 shows the variation of the heartbeat in each of the rounds of each participant. It is possible to notice that almost all of the variations were negative, meaning

TABLE 1.46 – ECG average BPM felled by the participants.

Participant	Visual Condition	Round	Baseline	Base	Audio	Haptic Belt	Virtual Cane	Mixture
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

TABLE 1.47 – ECG average BPM variation between rounds.

Participant	Base	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	-5.2%	2.7%	-2.9%	3.0%	12.7%	Sight
001C	-6.2%	-3.5%	-6.9%	8.7%	3.7%	Blind
002C	7.7%	23.0%	21.0%	21.0%	11.5%	Blind
003	-7.6%	14.6%	-0.8%	-4.8%	0.3%	Sight
003C	-1.5%	-3.5%	-1.8%	-0.8%	0.6%	Blind
004	5.3%	3.9%	5.1%	-2.1%	5.9%	Sight
004C	-0.5%	1.7%	0.4%	1.0%	-9.0%	Blind
005	-3.6%	-2.5%	-1.4%	0.8%	-2.4%	Sight

that the user decreased its workload between the "Baseline" and each method.

The Shapiro–Wilk normality test on the Table 1.49 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 1.50 there is no difference in the heart rate frequency variation between the sample groups.

The Table 1.51 shows the Anova test p-value of the heart rate frequency of the "blind"

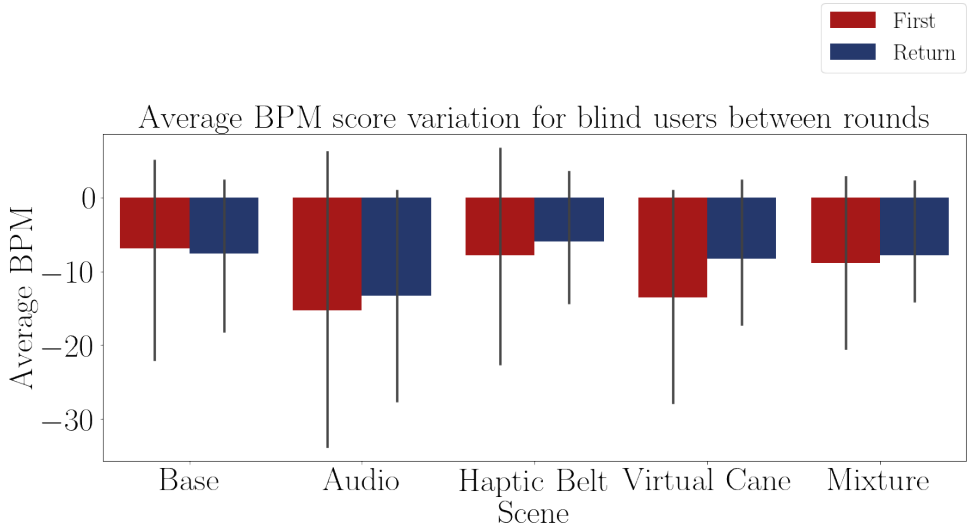

FIGURE 1.23 – Bar plot of the average heart rate of the blind participants on each method.

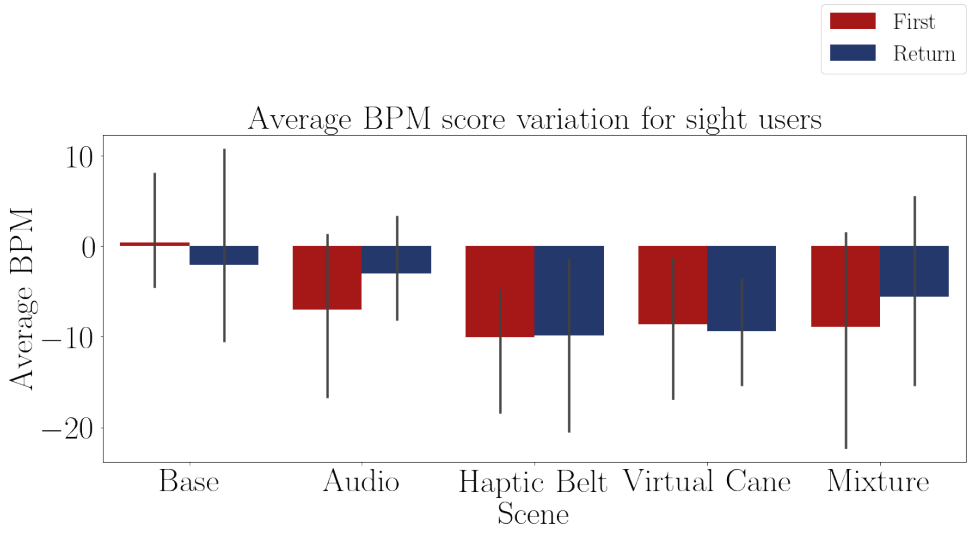

FIGURE 1.24 – Bar plot of the average heart rate of the sighted participants on each method.

sample between the guidance methods presented in the Table 1.48. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.

The Table 1.52 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 only "Audio" caused a different variation than the one noticed on the "Base" Method.

According to the Anova test at Table 1.51 and the LSD test at 1.52 only the "Audio" method provoked a different reaction than the "Base" method. Besides that, all methods decreased their frequency, going against the original expectations.

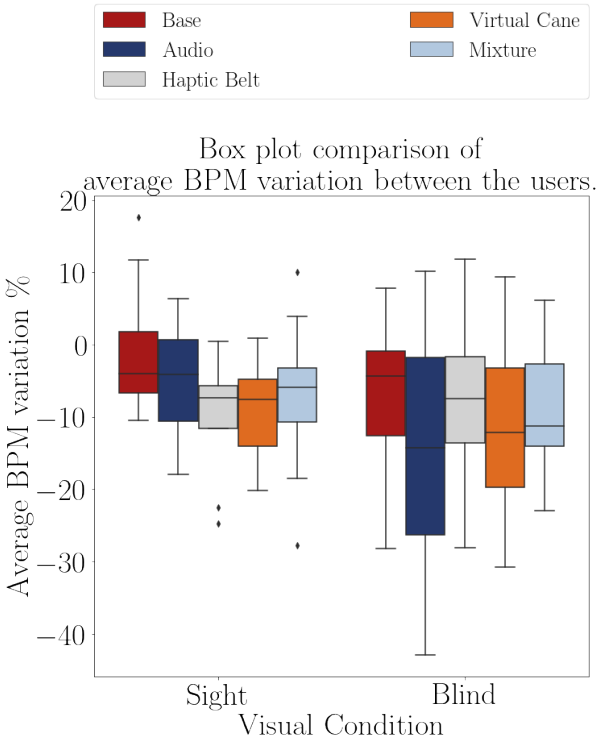

FIGURE 1.25 – Boxplot of the average heart rate of participants on each method.

TABLE 1.48 – ECG average BPM variation in relation to the baseline by participant and method.

Participant	Visual Condition	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	-5.4%	-12.4%	-22.5%	-20.2%	-27.7%
		Return	-10.3%	-10.0%	-24.7%	-17.9%	-18.5%
001C	Blind	First	-3.3%	-22.5%	-9.1%	-24.6%	-12.9%
		Return	-9.3%	-25.2%	-15.5%	-18.0%	-9.7%
002C	Blind	First	-28.2%	-42.9%	-28.1%	-30.8%	-22.9%
		Return	-22.6%	-29.8%	-13.0%	-16.3%	-14.1%
003	Sight	First	-3.1%	-18.0%	-7.2%	-8.4%	-6.0%
		Return	-10.5%	-6.0%	-8.0%	-12.8%	-5.6%
003C	Blind	First	7.8%	10.2%	11.8%	9.4%	5.5%
		Return	6.1%	6.3%	9.8%	8.5%	6.2%
004	Sight	First	11.7%	2.3%	-4.4%	0.9%	3.9%
		Return	17.6%	6.4%	0.5%	-1.1%	10.0%
004C	Blind	First	-4.1%	-6.1%	-5.9%	-8.1%	-5.5%
		Return	-4.5%	-4.5%	-5.5%	-7.2%	-14.0%
005	Sight	First	-1.5%	0.1%	-6.1%	-6.7%	-5.9%
		Return	-5.0%	-2.4%	-7.4%	-6.0%	-8.1%

TABLE 1.49 – Shapiro test p-value for the ecg average BPM for each method and visual condition

Method	Shapiro P-Value
Base blind	0.377
Base sight	0.086
Audio blind	0.721
Audio sight	0.969
Haptic Belt blind	0.665
Haptic Belt sight	0.059
Virtual Cane blind	0.584
Virtual Cane sight	0.743
Mixture blind	0.379
Mixture sight	0.663

TABLE 1.50 – T test p-value for the ecg average BPM each method for blinded users versus sighted users.

Method	T-Test P-Value
Base	0.279
Audio	0.215
Haptic Belt	0.594
Virtual Cane	0.750
Mixture	0.834

TABLE 1.51 – Anova p-value for the average BPM on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Between factors	303.579	4	75.895	0.383	0.819
Inside factors	6928.578	35	197.959		
Total	7232.157	39			

TABLE 1.52 – Cross validation p-value for the average BPM on each method for blinded users.

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

1.3.1.2 Analysis of the heartbeat frequency

The Table 1.53 presents the standard deviation of the interbeat interval by each participant on each scenes and they are plotted in the Figures 1.26 and 1.27. It is possible to see that there were all of the users felt an increase in the heartbeat variance.

TABLE 1.53 – ECG Average SDNN felled by the participants.

Participant	Visual Condition	Round	Baseline	Base	Audio	Haptic Belt	Virtual Cane	Mixture
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

The Figure 1.25 shows a comparison between both groups

The Table 1.54 shows the variation of the heartbeat in each round of each participant. In general, all the standard deviations increased, meaning that the mental workload decreased between the "Baseline" and the method.

The Shapiro–Wilk normality test on the Table 1.55 shows that all of the "blind" sample data are normally distributed, except the "Mixture" method. In the "sight" sample only the "Base" and the "Audio" method are normally distributed. That means that the following analyses cannot be made with those exceptions.

According to the T-Test presented in the Table 1.50 there is no difference in the heart rate frequency variation between the sample groups.

The Table 1.51 shows the Anova test p-value of the heart rate frequency of the "blind" sample between the guidance methods presented in the Table 1.48. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.

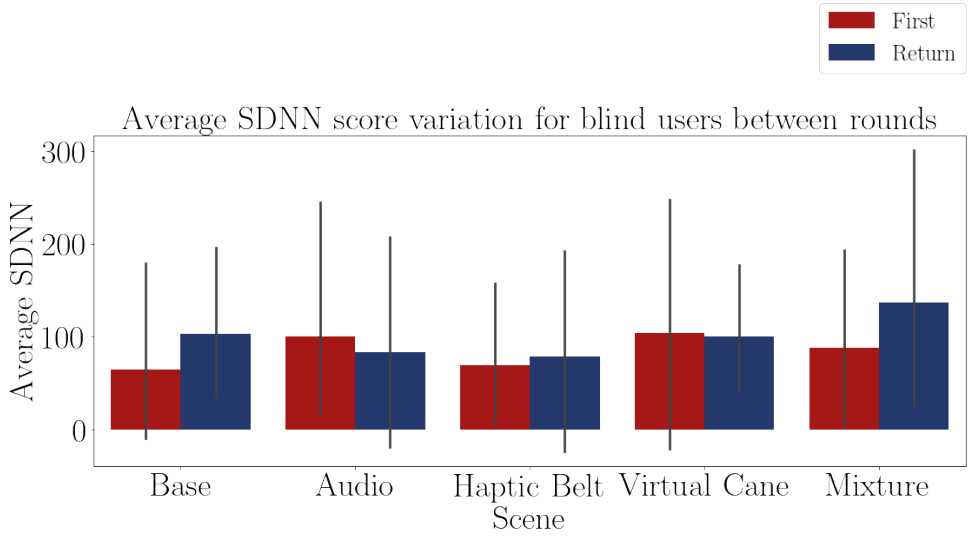

FIGURE 1.26 – Bar plot of the standard deviation of the heart of the blind participants on each method.

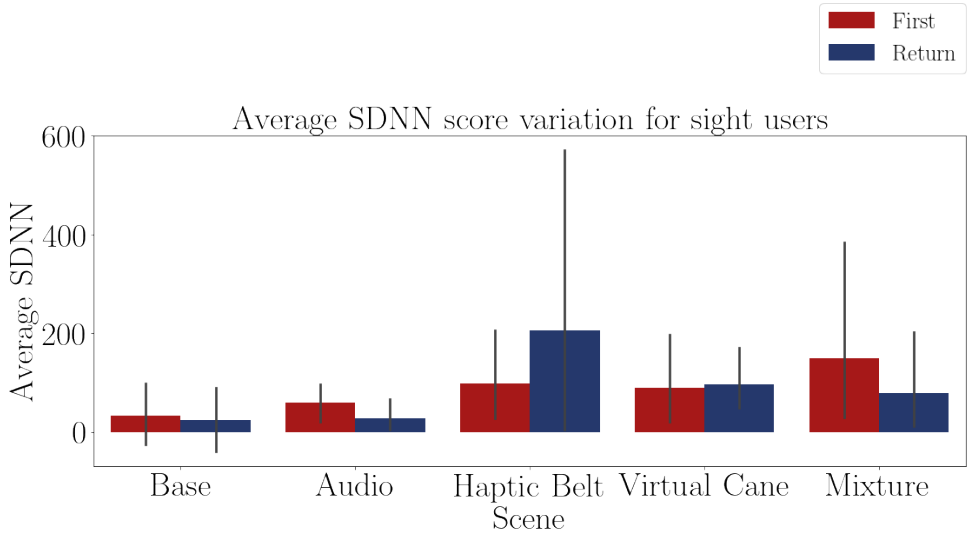

FIGURE 1.27 – Bar plot of the standard deviation of the heart of the sighted participants on each method.

The Table 1.52 presents the conclusion of a pairwise Fisher LSD test of the blind heart rate frequency variation between all the guidance methods. The results show no difference between the methods.

According to the Anova test at Table 1.57 and the LSD test at 1.58 there are no differences between the methods. Besides that, all of them provoked a decrease in the mental workload.

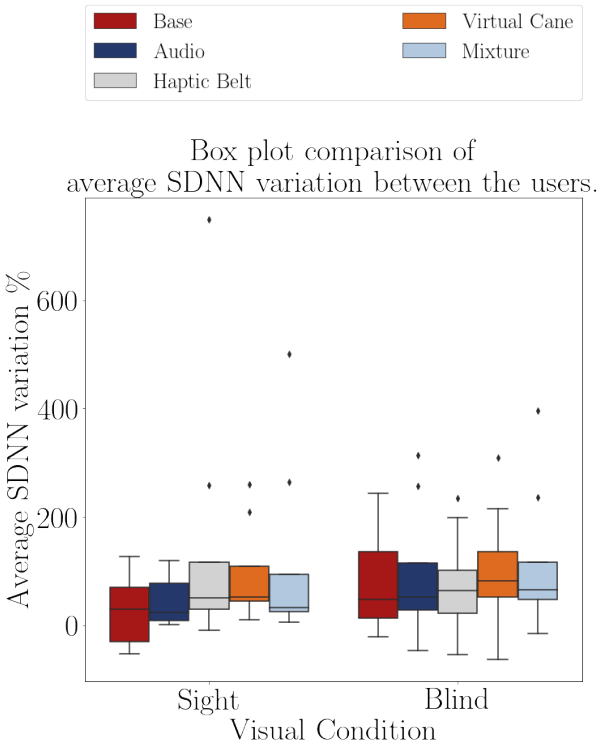

FIGURE 1.28 – Boxplot of the average heart rate of the participants on each method.

TABLE 1.54 – ECG Average SDNN variation in relation to the baseline by participant and method.

Participant	Visual Condition	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	120.5%	119.0%	258.5%	259.2%	500.7%
		Return	126.4%	85.2%	749.4%	209.1%	263.8%
001C	Blind	First	3.5%	36.3%	58.8%	108.7%	64.3%
		Return	53.7%	66.6%	67.5%	100.6%	58.9%
002C	Blind	First	-21.3%	5.4%	-13.5%	-63.8%	-15.4%
		Return	16.2%	-47.1%	-54.3%	21.6%	14.7%
003	Sight	First	27.9%	75.3%	14.1%	51.3%	34.0%
		Return	-53.1%	0.7%	-9.8%	46.1%	5.3%
003C	Blind	First	41.0%	46.6%	34.3%	62.2%	67.2%
		Return	101.8%	57.6%	69.3%	63.0%	76.5%
004	Sight	First	31.3%	32.0%	68.6%	40.0%	37.2%
		Return	52.4%	9.3%	33.5%	52.7%	29.9%
004C	Blind	First	237.1%	313.9%	198.2%	309.5%	235.9%
		Return	243.0%	257.0%	233.8%	215.0%	395.9%
005	Sight	First	-44.8%	8.8%	49.5%	9.9%	27.5%
		Return	-25.8%	15.6%	52.4%	75.3%	19.1%

TABLE 1.55 – Shapiro test p-value for the ecg average SDNN for each method and visual condition

Method	Shapiro P-Value
Base blind	0.078
Base sight	0.347
Audio blind	0.071
Audio sight	0.130
Haptic Belt blind	0.414
Haptic Belt sight	0.001
Virtual Cane blind	0.723
Virtual Cane sight	0.015
Mixture blind	0.027
Mixture sight	0.001

TABLE 1.56 – T test p-value for the ecg average SDNN each method for blinded users versus sighted users.

Method	T-Test P-Value
Base	0.230
Audio	0.317
Haptic Belt	0.434
Virtual Cane	0.862
Mixture	0.976

TABLE 1.57 – Anova p-value for the average SDNN on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Between factors	7039.359	4	1759.840	0.130	0.970
Inside factors	474190.070	35	13548.288		
Total	481229.429	39			

TABLE 1.58 – Cross validation p-value for the average SDNN 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	$H_0 : \mu_{Base} = \mu_{VirtualCane}$
Base	X	Mixture	$H_0 : \mu_{Base} = \mu_{Mixture}$
Audio	X	Haptic Belt	$H_0 : \mu_{Audio} = \mu_{HapticBelt}$
Audio	X	Virtual Cane	$H_0 : \mu_{Audio} = \mu_{VirtualCane}$
Audio	X	Mixture	$H_0 : \mu_{Audio} = \mu_{Mixture}$
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	$H_0 : \mu_{VirtualCane} = \mu_{Mixture}$

1.3.2 Galvanic skin reaction and temperature data;

The GSR analysis is made by analyzing the average in each round and comparing it with the "Baseline" average. The temperature was analyzed with the GSR to see if there is some influence and by a graphical analysis there was none.

The Table 1.59 presents the average skin conductance by each participant on each scenes and they are plotted in the Figures 1.29 and 1.30. 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.

TABLE 1.59 – GSR Average felled by the participants.

Participant	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	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
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

The Figure 1.25 shows a comparison between both groups

The Table 1.60 shows the variation of the heartbeat in each round of each participant. It is also possible to notice the same increase noticed before.

The Shapiro–Wilk normality test on the Table 1.61 shows that only the "Audio" method is normally distributed for the "blind" sample while for the "sight" sample only the "Virtual Cane" is not normally distributed

According to the T-Test presented in the Table 1.62 there is no difference in the heart rate frequency variation between the sample groups.

The Table 1.63 shows the Anova test p-value of the heart rate frequency of the "blind"

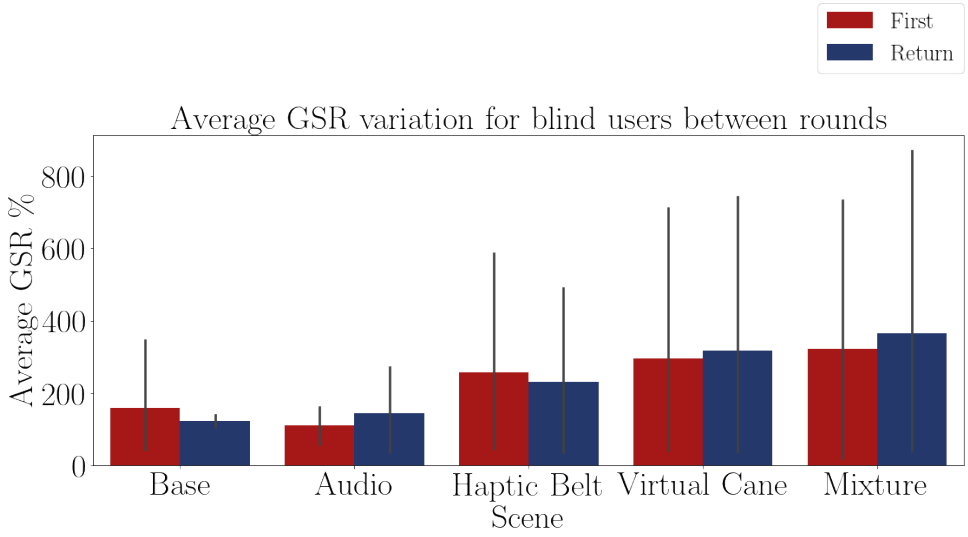

FIGURE 1.29 – Bar plot of the average skin conductance of the blind participants on each method.

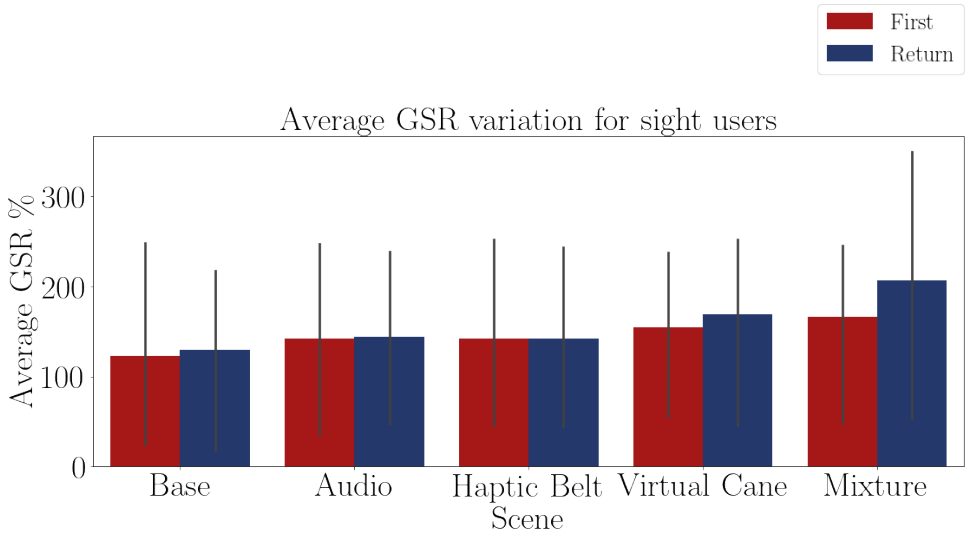

FIGURE 1.30 – Bar plot of the average skin conductance of the sighted participants on each method.

sample between the guidance methods presented in the Table 1.60. The p-value indicates that there is at least one method that is statistically equal to one of the other methods.

The Table 1.64 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" have different variations, but since they are not normally distributed this conclusion can not statistically be made.

According to the Anova test at Table 1.63 and the LSD test at 1.64 only the "Virtual Cane" and the "Mixture" method provoked a different reaction than the "Base" method, but since the Shapiro test at the Table 1.61 showed that they are not normally distributed, than this conclusion has no foundation.

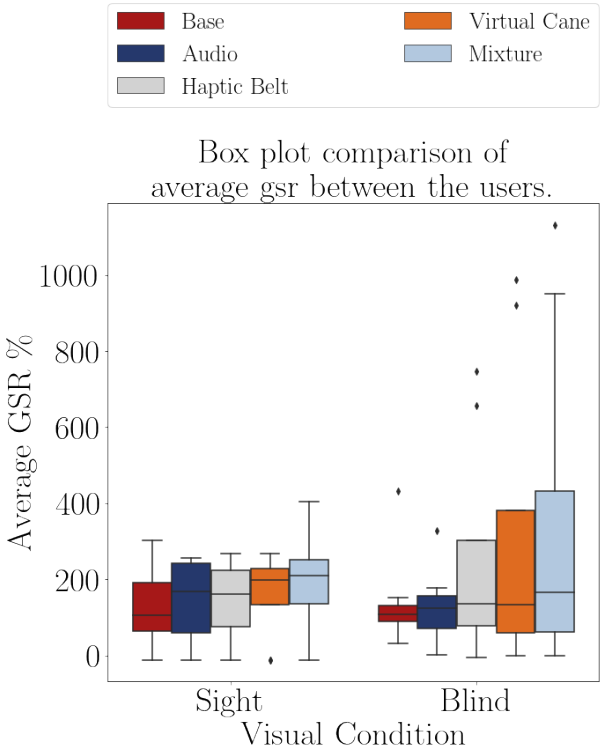

FIGURE 1.31 – Boxplot of the average skin conductace of the participants on each method.

TABLE 1.60 – GSR average variation in relation to the baseline by participant and method.

Participant	Visual Condition	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	106.1%	255.8%	266.9%	255.7%	231.5%
		Return	168.9%	250.2%	253.3%	268.2%	403.9%
001C	Blind	First	30.6%	176.5%	746.1%	920.7%	951.7%
		Return	125.3%	327.4%	657.0%	988.9%	1132.4%
002C	Blind	First	432.7%	32.3%	-0.0%	0.0%	0.0%
		Return	151.7%	1.7%	-5.1%	0.0%	0.0%
003	Sight	First	-3.0%	-12.0%	-12.0%	-12.0%	-11.9%
		Return	-12.0%	-12.0%	-11.9%	-12.0%	-11.9%
003C	Blind	First	85.4%	84.2%	104.2%	182.4%	258.8%
		Return	105.3%	109.2%	113.0%	202.4%	249.7%
004	Sight	First	85.4%	84.2%	104.2%	182.4%	258.8%
		Return	105.3%	109.2%	113.0%	202.4%	249.7%
004C	Blind	First	89.6%	148.5%	182.8%	84.3%	80.7%
		Return	108.2%	138.6%	159.0%	78.7%	81.6%
005	Sight	First	302.5%	239.2%	207.7%	193.9%	184.7%
		Return	255.2%	227.1%	214.9%	219.6%	185.9%

TABLE 1.61 – Shapiro test p-value for the gsr average for each method and visual condition

Method	Shapiro P-Value
Base blind	0.002
Base sight	0.565
Audio blind	0.544
Audio sight	0.065
Haptic Belt blind	0.017
Haptic Belt sight	0.194
Virtual Cane blind	0.004
Virtual Cane sight	0.020
Mixture blind	0.011
Mixture sight	0.281

TABLE 1.62 – T test p-value for the average GSR on each method for blinded users versus sighted users.

Method	T-Test P-Value
Base	0.802
Audio	0.780
Haptic Belt	0.367
Virtual Cane	0.348
Mixture	0.354

TABLE 1.63 – Anova p-value for the average GSR on each method for blinded users.

Source	Squared sum	DOF	Squared average	F	P-Value ($F_0 > F$)
Between factors	301240.786	4	75310.197	0.797	0.535
Inside factors	3307916.688	35	94511.905		
Total	3609157.475	39			

TABLE 1.64 – Cross validation p-value for the average 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	$H_0 : \mu_{Audio} = \mu_{HapticBelt}$
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	$H_0 : \mu_{VirtualCane} = \mu_{Mixture}$

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.

2.1 Do BVI users feel present in the VE as if they were in the real world?

2.1.1 Answers based on the simulation data

Analyzing the time that each user took to complete each scene, it is not possible to infer a conclusion, because this data was not meant to measure this goal.

2.1.2 Answers based on the subjective data

This data also was not made to assess this goal, because there is no base of comparison with subjective data from before the experiment, hence before the user started to use the virtual reality.

2.1.3 Answers based on the physiological data

According to the ECG data, there was a decrease in the mental workload during the experiment while the expectation was to be an increase instead. This difference proves that the users were impacted by the experiment in the virtual reality, but does not represent the same situation outside the virtual reality.

The GSR data also showed a change when the users were using virtual reality. The results showed that the users were aroused or had an increased mental workload.

2.1.4 Final conclusions

The physiological data gathered was the only source of data to assess this goal, and they had opposite conclusions regarding the expected. The ECG showed a decrease in the mental workload, while the GSR showed an increase.

The collected ECG data is less reliable than the collected GSR due to the sensibility of the sensors used. It was noted that the ECG is very sensitive to movements and the position of the sensors in relationship with the data receiver. If something stands in the way between the sensor and receiver, such as a human body, that data is lost, causing the resulting analysis to be noiseless or to be made using corrections such as the one used.

So, this goal was partially achieved.

2.2 Does BVI users rely more on one type of information than the other?

2.2.1 Answers based on the simulation data

With the regard to the time, and since the Anova test showed that all of the data are different from each other, one can look at the Table 1.4 and notice that the method that the users took the shorter time was the ”Mixture” method, which was along with the expectation that the BVI users rely on both of the information, but the second shortest was the ”Virtual Cane”, so it indicates that the BVI user relies more on a haptic source of information. But this data is not entirely reliable, since there were a couple of mistakes in the first experiment to close each simulation, hence increasing the final time of each user at the round.

2.2.2 Answers based on the subjective data

Analysing the Figures 1.7, 1.12, 1.17 and 1.22 one can notice that the haptic source of information is preferable for they have the best results in general for each questionnaire, but the Anova test disagree with that conclusion in some cases, but that can a conseguence of the fact that only 4 individuals of each group did the experiment.

2.2.3 Answers based on the physiological data

Disconsidering that all of the ECG data were against the expected variation, according to the Anova test, only the ”Audio” method can be concluded that is different from the

”Base” method and Figure 1.23 shows a similar conclusion.

The skin conductance Anova test resulted that only the ”Virtual Cane” and the ”Mixture” method are different than the ”Base” method, and, also according to the skin data, they aroused more the user or have a higher mental workload. Another conclusion from the Anova test is that the ”Audio” method has a similar workload to the ”Base” method, and ironically this was the only one that could be said that arouses less or has a lesser mental workload.

2.2.4 Final conclusions and comentaries

The majority of the graphics showed a tendency that haptics sources of data are more favorable for the BVI users, but the conclusion drawn by the hypothesis test did not support that analysis. This may be due to the small sample size.

One observation made during the experiment is that the BVI users during the ”Audio” and ”Mixture” method did not use, or used only a few times, the audio guidance provided by the researcher. This does not discard that they did not rely on sound information, because the simulation has filled with audio cues. This may be because of their previous experience in navigation and mobility alone.

The conclusion for this goal would be that they do rely more on a mixture of haptic and audio data, the first for obstacle detection at and short distance, the latter for guidance and information gathered at bigger distances.

2.3 Do non-BVI users have the same demands and skills as BVI users when designing assistive products?

2.3.1 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.

Results from the Figure 1.3 showed 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.

2.3.2 Answers based on the subjective data

The T-Test of each questionnaire showed that there are no differences between the groups, but the Figures 1.7, 1.12 and 1.17 showed a noticeable difference between the groups. These unmatched results may be because of the small sample number.

2.3.3 Answers based on the physiological data

The Figure 1.28 indicates that there are no visual difference between the groups. The Figure 1.25 indicates a difference on the distribution and a rather similar average. The Figure 1.31 indicates higher arousal or mental workload by the ”blind” users. All the T-Tests indicate that both groups have the same variation of workload and arouse.

2.3.4 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.

To close up, this goal is considered not achieved for lack of a bigger and more diverse sample size.

2.4 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;
- Repeat the experiment with more devices with different proportions of haptic and audio information sources;
- 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;