Analysis of the effects between Independent Variable and Dependent Variable

Statistical Test Methods

The statistical test methods are listed in the table below. The motion states of the visual distractions are a categorical and nominal variable. The maze completion time, measured in seconds, should be considered as a continuous ratio variable. Thus the statistic test method we choose is t-test. The second dependent variable is the landmarks recall test score, which has five possible values: 0, 1, 2, 3 and 4. It is considered as a categorical variable and thus the statistic test method we choose is Chi-squared test. The third dependent variable is the point-to-the-end test score, which has two possible values: right and wrong. It is considered as a categorical variable and thus the statistic test method we choose is Chi-squared test.

Table 2. Statistical Test Methods

	DV1: Maze Completion Time (Ratio+Continuous)	DV2: Landmarks Recall Test score (Categorical)	DV3: Point-to-the-end test score (Categorical)
IV: Dynamic / Static Visual Distractions (Categorical)	t-test	Chi-squared test	Chi-squared test

Statistical Test Results

With α=0.05, there is no significant difference between the impact of dynamic and static distractions on navigation performance. Nor is there such difference between the impact of dynamic and static distractions on spatial memory performance, or between the impact of dynamic and static distractions on visual memory performance. However, with α=0.1, there is a significant difference between the impact of dynamic and static distractions on visual memory performance.

In a nutshell, we do not find a statistical difference between our two groups, and therefore we fail to reject the null hypothesis. However, with α=0.1, such difference exists between the impact of dynamic and static distractions on visual memory performance.

Table 3. Statistical Test Results

	IV~DV1: Maze Completion Time (Ratio+Continuous)	IV~DV2: Landmarks Recall Test Score (Categorical)	IV~DV3: Point-to-the-end Test Score (Categorical)
Method	Unpaired two-samples t-test	Chi-squared test	Chi-squared test
p-value	0.461	0.08894	0.8087
Significant?	N	N	N

Statistical Test Details

The Effects of Dynamic and Static Visual Distractions on Maze Completion Time

Data Exploration and Visualization

Table 4. Data Exploration of the Maze Completion Time by groups

Group	Count	Mean	sd
1	9	99.78	55.83
2	7	77.00	64.34

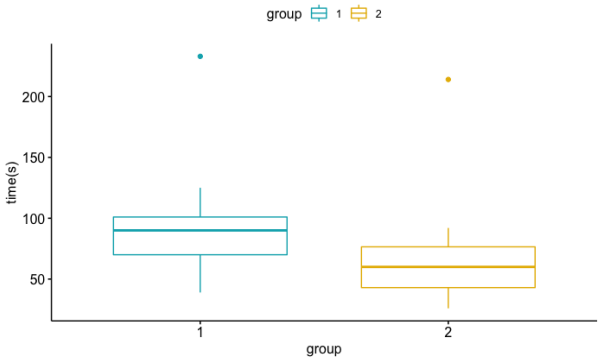

Figure 3. Boxplots for groups

Assumption Examination

First, the two groups are independent since the samples from group 1 and group 2 are not related.
Second, we used the Shapiro-Wilk normality test to find out whether the data from the two groups is normally distributed. The null hypothesis of the test is that the data is normally distributed, while the alternative hypothesis being the data is not normally distributed. We use functions with() and shapiro.test() to conduct the Shapiro-Wilk test for each group of samples.

Code and result:

```
with(data, shapiro.test(time[group == "1"]))# p =0.28
with(data, shapiro.test(time[group == "2"]))# p =0.13
```

From the output, we discover that the two p-values are greater than the significance level 0.05 implying that the distribution of the data is not significantly different from the normal distribution. In other words, we can assume normality.
Third, we investigate that if the two populations have the same variances. We use F-test to test for homogeneity in variances. This is performed with the function var.test() as follows:

Code:

```
var.test(time ~ group, data = data)
```

Result:

```
F-test to compare two variances
data:  time by group
F = 0.75282, num df = 8, denom df = 6, p-value = 0.691
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.134442 3.501909
sample estimates:
ratio of variances
         0.7528243
```

The p-value of F-test is p = 0.691. It's greater than the significance level of α = 0.05. In conclusion, there is no significant difference between the variances of the two sets of data. Therefore, we can use the classic t-test which assumes equality of the two variances.

Compute unpaired two-samples t-test

Code:

```
t.test(time ~ group, data = data, var.equal = TRUE)
```

Result:

```
Two Sample t-test

data:  time by group
t = 0.75797, df = 14, p-value = 0.461
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -41.67536  87.23091
sample estimates:
mean in group 1 mean in group 2
       99.77778        77.00000
```

The p-value of the test is 0.461, which is greater than the significance level α = 0.05. We can conclude that group 1's maze completion time is not significantly different from group 2's maze completion time.

The Effects of Dynamic/Static Visual Distractions on Landmarks Recall Test Score

Data Exploration and Visualization

Table 5. Data Exploration of the Landmarks Recall Test Score by groups

	Group 1	Group 2
1 correct answer	3	3
2 correct answers	5	1
3 correct answers	0	3
4 correct answers	1	0

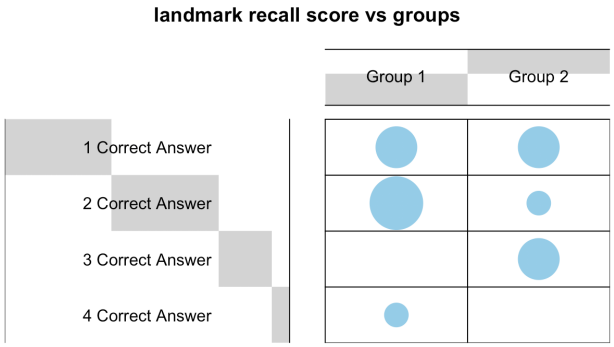

Figure 4. Landmark result-test score distribution

Hypothesis

Chi-square test examines whether rows and columns of a contingency table are statistically significantly associated. The null hypothesis (H0) is that the row and column variables of the contingency table are independent while the alternative hypothesis (H1) being the row and column variables are dependent.

Code and Result

```
tbl = table(data3$score_land, data3$group)
colnames(tbl) <- c("Group 1", "Group 2")
rownames(tbl) <- c("1 Correct Answer", "2 Correct Answer", "3 Correct Answer", "4
Correct Answer")
chisq.test(tbl)

Pearson's Chi-squared test
data:  tbl
X-squared = 6.5185, df = 3, p-value = 0.08894
```

Conclusion

From the output, the row and the column variables are not statistically significantly associated (p-value = 0.08894).

The Effects of Dynamic/Static Visual Distractions on Point-to-the-end Test Score

Data Exploration and Visualization

Table 6. Data Exploration of the Point-to-the-end Test Score by groups

	Group 1	Group 2

correct	8	5
wrong	1	2

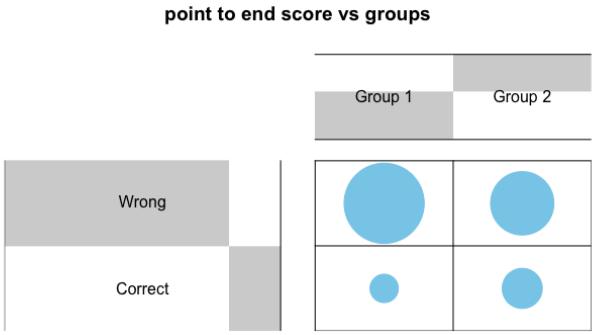

Figure 5. Point-to-the-end score distribution

Hypothesis

Chi-square test examines whether rows and columns of a contingency table are statistically significantly associated. The null hypothesis (H0) is that the row and column variables of the contingency table are independent and the alternative hypothesis (H1) is that the row and column variables are dependent.

Code and Result

```
tbl = table(data3$score_point, data3$group)
colnames(tbl) <- c("Group 1", "Group 2")
rownames(tbl) <- c("Wrong", "Correct")
chisq.test(tbl)

Pearson's Chi-squared test with Yates' continuity correction

data:  tbl
X-squared = 0.058608, df = 1, p-value = 0.8087
```

Conclusion

From the output, the row and the column variables are not statistically significantly associated (p-value = 0.8087).

Statistical Test of Potential Confounding Effects

Statistical Test Methods

We then explored the impact of potential confounding variables on navigation performance, spatial memory performance, and visual memory performance. We collected the following as potential confounding variables: spatial memory test score, visual memory test score, gender, previous video game experience, previous VR experience, and previous Minecraft experience.

Table 7. Statistical Test Methods of the Potential Confounding Effects

	V1: Maze Completion Time (Ratio+Continuous)	V2: Landmarks Recall Test Score (Categorical)	V3: Point-to-the-end Test Score (Categorical)
Spatial Memory Test Score (Ratio+Continuous)	Pearson's correlation test	Kruskal-Wallis one-way ANOVA test	t-test
Visual Memory Test Score (Ratio+Continuous)	Pearson's correlation test	Kruskal-Wallis one-way ANOVA test	t-test
Gender (Categorical)	t-test	Chi-squared test	Chi-squared test
Previous Video Game Experience(Categorical)	t-test	Chi-squared test	Chi-squared test
Previous VR Experience(Categorical)	t-test	Chi-squared test	Chi-squared test
Previous Minecraft Experience(Categorical)	t-test	Chi-squared test	Chi-squared test

Statistical Test Result

Table 8. Statistical Test Results(p-value) of the Potential Confounding Effects

	V1: Maze Completion Time (Ratio+Continuous)	V2: Landmarks Recall Test Score (Categorical)	V3: Point-to-the-end Test Score (Categorical)
Spatial Memory Test Score (Ratio+Continuous)	0.0225	0.7651	0.0918
Visual Memory Test Score (Ratio+Continuous)	0.0727	0.9435	0.1775
Gender (Categorical)	0.9224	0.6198	0.2942
Previous Video Game Experience(Categorical)	0.764	0.6198	0.0014
Previous VR Experience(Categorical)	0.9034	0.7851	0.0688
Previous Minecraft Experience(Categorical)	0.7379	0.6198	0.9183

With α=0.05, we found two potential confounding effects: spatial memory test score may have a confounding effect on navigation performance; previous video game experience is highly likely to have a confounding effect on spatial memory performance. With α=0.1, we found several additional potential confounding effect: previous VR experience may have a confounding effect on spatial memory performance; visual memory test score may have a confounding effect on navigation performance; spatial memory test score may have a confounding effect on spatial memory performance

Power Analysis

In each group, to obtain a power of 0.80, when the effect size is moderate (0.3) and a significance level of 0.05 is employed, the corresponding recommended sample sizes are listed in the table below

Table 9. Power Analysis

	DV1: Maze Completion Time	DV2: Landmarks Recall Test Score	DV3: Point-to-the-end Test Score
IV: Dynamic/Static Visual Distractions	96	44	121