

Contents

1	Introduction	2
2	Design Evaluation Experiment	3
2.1	Introduction	3
2.2	Experimental Design	4
2.2.1	Task Design	4
2.2.2	Instrument Design	4
2.3	Methods	4
2.4	Results	5
2.4.1	Demographics	5
2.4.2	Statistical Tests	5
2.4.3	Performance Measures	6
2.4.4	Design Feedback	8
2.5	Discussion	8
2.6	Conclusion	8
Appendices		9

¹⁷ Chapter 1

¹⁸ Introduction

Chapter 2

Design Evaluation Experiment

2.1 Introduction

After investigating the technical approach and the benefit to including the passive haptics layer, we seek to investigate the use of the Rapidly Reconfigurable Research Cockpit in a more realistic design evaluation study. The advantages of using the R3C system would not be useful if it masked defects in a design study.

2.2 Experimental Design

2.2.1 Task Design

The task the subjects were to perform had a number of requirements.

- Ability to simulate designs for completing task on touchscreen and R3C setup
- Tracking task using a standard attitude indicator display controlled with joystick
- Second task that requires use of multiple button to button movements on the instrument
- Sufficient workload such that subjects have high but not full workload

2.2.2 Instrument Design

2.3 Methods

Subjects were divided into the two groups, TS and VR.

2.4 Results

2.4.1 Demographics

Twenty-three subjects were recruited from the UC Davis engineering undergraduate and graduate student population. Twelve subjects were placed in the VR group, and the remaining eleven in the TS group. The mean age was $21.0(\sigma = 3.14)$, with 19 male and 4 female subjects. The female subjects were balanced between the two groups. Most subjects had no flight experience (two were student pilots), and all of the VR group subjects indicated that they had less than one hour of experience using virtual reality headsets.

2.4.2 Statistical Tests

The quantitative dependent measures are tested with a two-way ANOVA, with one within subjects factor (Design) and one between subjects factor (Group). The Design factor contains two levels, the two designs each subject tested, Edgekey and Keypad. The Group factor also contains two levels, the VR group and the TS group. When the ANOVA showed significance in the interaction test, post-hoc repeated measured t-tests were un-

dertaken to determine the significance of Design within each Group. All effects were considered statistically significant at the 0.0125 level. Statistical significance level was corrected using the Bonferroni correction considering the 4 different dependent measures being tested ($\alpha = 0.05/4 = 0.0125$).

2.4.3 Performance Measures

The performance of the tracking task was measured using the root-mean square average (RMSE) of the error. The effect of group yielded an F ratio of $F(1, 21) = 21.4, p < 0.001$ indicating a significant difference between VR ($M = 1.28\text{deg}, \sigma = 0.38\text{deg}$) and TS ($M = 1.97\text{deg}, \sigma = 0.38\text{deg}$). The effect of design indicated no significant difference ($F(1, 21) = 5.94, p = 0.024$) between Keypad ($M = 1.57\text{deg}, \sigma = 0.51\text{deg}$) and Edgekey ($M = 1.70\text{deg}, \sigma = 0.52\text{deg}$). The interaction effect was not significant ($F(1, 21) = 0.17, p = 0.69$).

Response time. The effect of group yielded an F ratio of $F(1, 21) = 1.61, p = 0.22$ indicating no significant difference between VR ($M = 2983\text{msec}, \sigma = 439\text{msec}$) and TS ($M = 2737\text{msec}, \sigma = 566\text{msec}$). The effect of design indicated a significant difference ($F(1, 21) = 13.9, p = 0.001$) between Keypad ($M = 2728\text{msec}, \sigma = 512\text{msec}$) and Edgekey ($M = 3002, \sigma = 488\text{msec}$).

The interaction effect was not significant ($F(1, 21) = 0.17, p = 0.69$).

Number of prompts correct. The effect of group yielded an F ratio of $F(1, 21) = 43.9, p < 0.001$ indicating a significant difference between VR ($M = 6.06, \sigma = 2.90$) and TS ($M = 10.2, \sigma = 1.23$). The effect of design indicated a significant difference ($F(1, 21) = 64.1, p < 0.001$) between Keypad ($M = 9.30, \sigma = 1.83$) and Edgekey ($M = 6.78, \sigma = 3.54$).

The interaction effect was significant as well ($F(1, 21) = 27.8, p < 0.001$).

The post hoc tests indicated significance between designs for the VR group ($t(11) = 8.0, p < 0.001$) between the Keypad design ($M = 8.11, \sigma = 1.62$) and the Edgekey ($M = 4.00, \sigma = 2.37$). The post hoc tests indicated no significant difference between designs for the TS group ($t(10) = 2.3, p = 0.05$) between the Keypad design ($M = 9.82, \sigma = 1.38$) and the Edgekey ($M = 10.6, \sigma = 0.96$).

NASA TLX scores. The effect of group yielded an F ratio of $F(1, 21) = 1.69, p = 0.21$ indicating a significant difference between VR ($M = 70.0, \sigma = 22.6$) and TS ($M = 65.3, \sigma = 8.53$). The effect of design indicated a significant difference ($F(1, 21) = 23.6, p < 0.001$) between Keypad ($M = 57.8, \sigma = 15.2$) and Edgekey ($M = 77.7, \sigma = 13.4$). The interaction effect was significant as well ($F(1, 21) = 8.25, p < 0.001$). The post hoc tests in-

94 dicated signifigance between designs for the VR group ($t(11) = -4.20, p =$
95 0.001) between the Keypad design ($M = 54.4, \sigma = 20.4$) and the Edgekey
96 ($M = 85.6, \sigma = 11.2$) The post hoc tests indicated no signifigant difference
97 between designs for the TS group ($t(10) = -2.72, p = 0.02$) between the
98 Keypad design ($M = 61.5, \sigma = 4.46$) and the Edgekey ($M = 69.2, \sigma =$
99 10.1)

100 **2.4.4 Design Feedback**

101 **2.5 Discussion**

102 **2.6 Conclusion**

Appendices