

# STA 322 Project 2

Alayna Binder & Katherine Zhong

2025-12-05

## 1 Experimental Design

We conducted a full  $2^4$  factorial experiment that included all sixteen combinations of the four helicopter design factors A (rotor length), B (leg length), C (leg width), and D (paper clip on leg). A full factorial structure was appropriate because our goal was to estimate the main effects of each design feature and evaluate whether any interactions influenced flight time, including interaction terms that might be aliased in a fractional factorial design. We collected six replicate drops for each design to obtain precise estimates of the factor effects while keeping the data collection workload manageable.

To reduce the possibility of ordering effects, we randomized the sequence of the sixteen design combinations before beginning data collection. After establishing this order, we also randomized which thrower completed each drop. Both throwers aimed to release the helicopters in a consistent manner, and random assignment ensures that any natural differences in release technique were distributed evenly across all treatments.

We treated thrower as a blocking variable. Although using a single thrower would theoretically eliminate person-to-person variation, it would also require one individual to perform all 96 drops. That workload increases the likelihood of fatigue or gradual shifts in release technique, which can introduce time-related variability that is difficult to control. We judged this within-person drift to be a greater risk than the small, consistent differences that might exist between two people. Because both project members were involved in data collection, blocking was the appropriate design choice. By structuring the experiment so that each thrower completed drops for every design combination, we created a fully crossed block that allowed us to account for person-specific differences directly and improve the precision of the estimated factor effects.

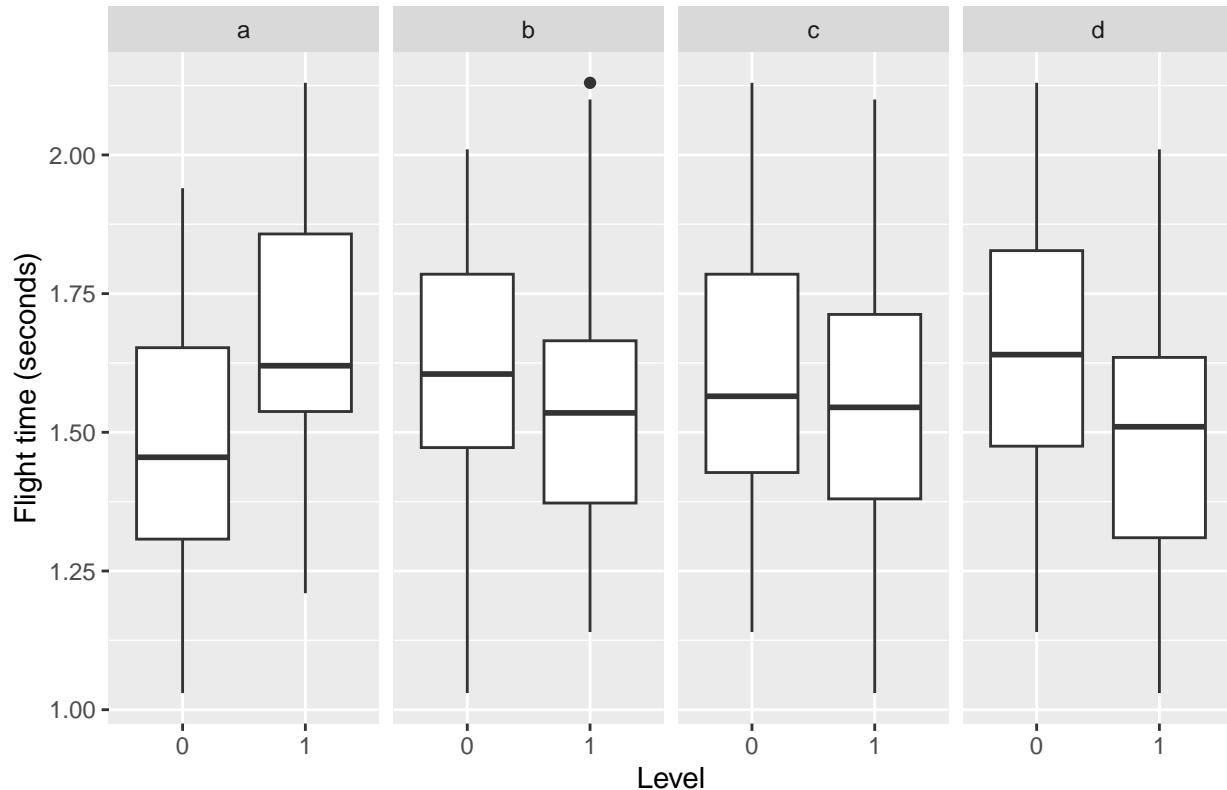
All drops were completed within a single session. While some within-session variation is possible, the number of drops per person was modest after splitting the workload, and the randomized allocation of treatments ensures that any minor temporal changes did not align with specific factor levels. We did not observe systematic shifts in flight time over the course of the session, so we treated within-session variation as negligible.

## 2 Analysis

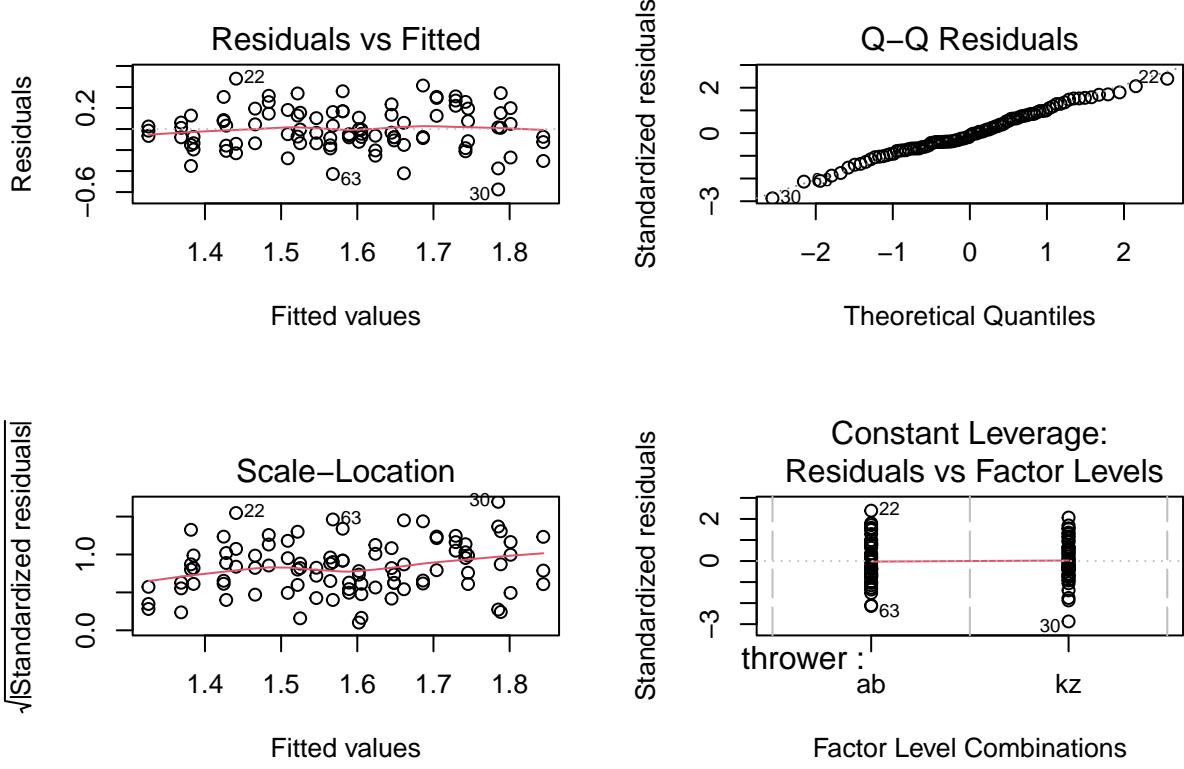
### 2.1 Question 1

To identify which design factors influence flight time, we first examined the distributions of flight times across the two levels of each factor. The main effect boxplots below show a clear upward shift for factor A, with level 1 producing noticeably longer flights than level 0. Factor D shows the opposite pattern, with shorter flights at level 1. In contrast, factors B and C display only small differences between levels, and the variation within each level is larger than any apparent shift between them.

## Main Effect Plots for Factors A–D



We then formalized these patterns using a linear model that included all four design factors while blocking on thrower. A check of the model assumptions shows that the diagnostics are acceptable: the residuals vs. fitted plot shows no systematic curvature, supporting the linearity assumption; the Q-Q plot follows the reference line closely aside from minor tail deviations, indicating approximate normality; and the scale-location plot suggests that the spread of residuals remains fairly constant across fitted values, indicating no strong heteroscedasticity. The constant leverage plot also shows no observations with both large residuals and high leverage, so there are no influential points that would compromise the model.



The estimated effects and confidence intervals are shown in the table below. Factor A had a positive and statistically significant effect, increasing the mean flight time by about 0.22 seconds. Factor D had a negative and significant effect, decreasing the mean flight time by about 0.14 seconds. The confidence intervals for both effects were entirely away from zero. Factors B and C had small estimated effects with confidence intervals that included zero, indicating little evidence that they meaningfully affect flight time.

Table 1: Main Effects Model Results

term	estimate	conf.low	conf.high	std.error	statistic	p.value
(Intercept)	1.624	1.521	1.726	0.052	31.411	0.000
a	0.220	0.137	0.304	0.042	5.222	0.000
b	-0.056	-0.140	0.028	0.042	-1.323	0.189
c	-0.043	-0.127	0.041	0.042	-1.017	0.312
d	-0.140	-0.224	-0.056	0.042	-3.317	0.001
throwerkz	-0.059	-0.143	0.025	0.042	-1.402	0.164

Both the exploratory plots and the fitted model point to the same conclusion. Factors A and D are the most influential design components. Factor A corresponds to rotor length, and the model indicates that longer rotors tend to increase flight time. Factor D corresponds to whether a paper clip is attached, and the negative effect estimate suggests that adding a paper clip shortens the flight. In contrast, the variables associated with leg length (factor B) and leg width (factor C) show minimal impact on flight duration under this experimental setup.

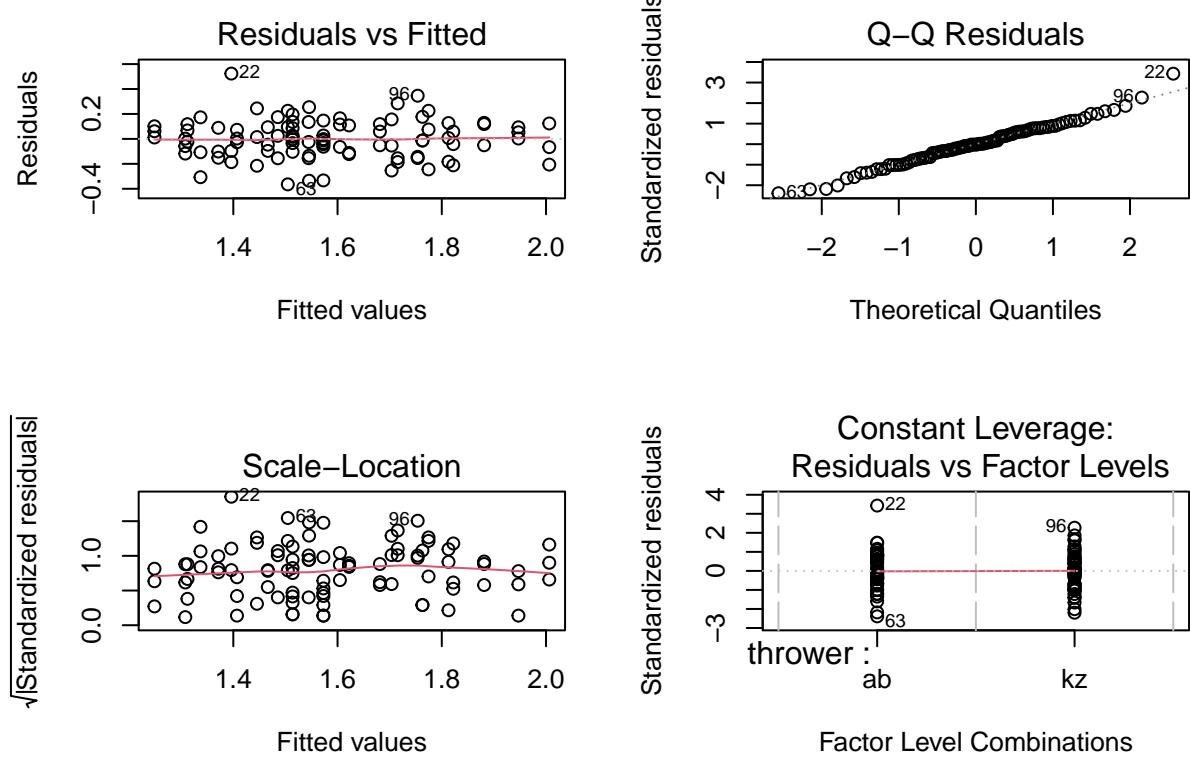
## 2.2 Question 2

We use a linear regression model to investigate this potential interaction effect between rotor length (A) and leg width (C).

First, we check the model assumptions of a full model with all main effects, all two-way, three-way and

four-way interactions between A, B, C, D, and `thrower` as the block variable. We do so by checking the residual vs. fitted plot, Q-Q plot, scale-location plot, and residuals vs. leverage plot.

We choose to have the block variable as a fixed effect in the model since we do not expect the thrower effect to be different based on the different helicopter features.



Overall, the model diagnostics suggest that the linear regression model is appropriate for these data. The residuals appear to be approximately normally distributed, and the residuals vs. fitted plot shows no clear patterns that would indicate non-linearity or heteroscedasticity. The Q-Q plot further supports the normality assumption since the residuals fall close to the diagonal reference line, meaning their distribution aligns well with what we would expect under normality, with no strong skewness or unusual tail behavior. The scale-location plot also shows a fairly horizontal trend, which indicates that the variance of the residuals remains roughly constant across levels of the fitted values. While a few potential outliers are present, the constant leverage plot suggests that they do not appear to have a substantial influence on the model or meaningfully compromise the overall fit.

Since we are interested in the two-way interaction between rotor length (A) and leg width (C), we fit a reduced model that includes only the main effects and all two-way interaction terms, along with the blocking variable `thrower`.

Then, we compare whether the model with only main effects and two-way interactions is significantly different from the full model with all interaction terms using an ANOVA test.

Table 2: ANOVA Comparison of Reduced and Full Models

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
84	2.912025	NA	NA	NA	NA
79	2.236583	5	0.6754417	4.771554	0.0007321

The ANOVA test comparing the full model with all interaction terms to the reduced model with only main

effects and two-way interactions yields a p-value of 0.0007. The p-value is less than 0.05, indicating that there is a significant difference between the two models, suggesting that the additional higher-order interaction terms in the full model does provide a better fit to the data.

Additionally, we fit a model with all main effects, two-way interaction effects, three-way interaction effects, and the blocking variable **thrower** to see if including the full model significantly improves the model fit compared to the three-way interaction model.

Table 3: ANOVA Comparison of Three-Way Interaction and Full Models

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
80	2.250600	NA	NA	NA	NA
79	2.236583	1	0.0140167	0.495093	0.4837329

The ANOVA test comparing the three-way interaction model to the full model yields a p-value of 0.4837. Since the p-value is greater than 0.05, this indicates that including the four-way interaction effect does not significantly improve the model fit compared to the three-way interaction model. Therefore, we will use the three-way interaction model for further analysis.

Lastly, for confirmation, we can also compare the two-way interaction model to the three-way interaction model using an ANOVA test.

Table 4: ANOVA Comparison of Two-Way and Three-Way Interaction Models

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
84	2.912025	NA	NA	NA	NA
80	2.250600	4	0.661425	5.877766	0.0003371

Since  $p\text{-value} = 0.0003 > 0.05$ , this indicates that including the three-way interactions model does significantly improve the model fit compared to the two-way interaction model.

Table 5: Linear Model with Three-Way Interaction Results

term	estimate	conf.low	conf.high	std.error	statistic	p.value
(Intercept)	1.454	1.318	1.590	0.068	21.237	0.000
a	0.163	-0.018	0.343	0.091	1.794	0.077
b	0.063	-0.118	0.243	0.091	0.690	0.492
c	0.321	0.141	0.501	0.091	3.542	0.001
d	0.239	0.059	0.419	0.091	2.640	0.010
throwerkz	-0.059	-0.127	0.009	0.034	-1.728	0.088
a:b	0.315	0.079	0.551	0.119	2.656	0.010
a:c	-0.175	-0.411	0.061	0.119	-1.476	0.144
a:d	0.013	-0.223	0.249	0.119	0.112	0.911
b:c	-0.305	-0.541	-0.069	0.119	-2.572	0.012
b:d	-0.397	-0.633	-0.161	0.119	-3.345	0.001
c:d	-0.630	-0.866	-0.394	0.119	-5.312	0.000
a:b:c	-0.010	-0.283	0.263	0.137	-0.073	0.942
a:b:d	-0.265	-0.538	0.008	0.137	-1.935	0.057
a:c:d	0.200	-0.073	0.473	0.137	1.460	0.148
b:c:d	0.575	0.302	0.848	0.137	4.199	0.000

Based on our regression model results above, the estimated interaction effect of AC is -0.175 with a p-value of 0.144. The estimated effect size is small and  $p > 0.05$ , indicating insufficient evidence to support the claim of a significant interaction effect between rotor length (A) and leg width (C) on flight time. The confidence interval for the AC interaction term also includes zero (-0.411, 0.061), further confirming that the interaction effect is not statistically significant.

### 2.3 Question 3

To determine the helicopter design that maximizes flight time, we estimated the average flight time for each of the 16 combinations of factors A-D using the three-way interaction model selected in Question 2 (as this model provided the best fit among all evaluated).

We obtained the estimated means using the `emmeans` package. We opted for this approach over manually resetting reference levels because it computes the model-based mean for every factor combination directly and consistently while properly incorporating all main effects, interactions, and the blocking variable.

Table 6: Estimated average flight time (seconds) for all 16 factor combinations

Levels (A,B,C,D)	Estimate (s)	95% CI low	95% CI high
(1, 1, 0, 0)	1.965	1.833	2.097
(1, 0, 0, 1)	1.840	1.708	1.972
(1, 1, 1, 0)	1.795	1.663	1.927
(0, 0, 1, 0)	1.745	1.613	1.877
(1, 0, 1, 0)	1.733	1.601	1.865
(0, 0, 0, 1)	1.664	1.532	1.796
(1, 0, 0, 0)	1.587	1.455	1.719
(1, 1, 0, 1)	1.555	1.423	1.687
(1, 0, 1, 1)	1.555	1.423	1.687
(1, 1, 1, 1)	1.531	1.399	1.663
(0, 1, 1, 0)	1.503	1.371	1.635
(0, 1, 0, 0)	1.487	1.355	1.619
(0, 0, 0, 0)	1.425	1.293	1.557
(0, 0, 1, 1)	1.355	1.223	1.487
(0, 1, 0, 1)	1.330	1.198	1.462
(0, 1, 1, 1)	1.290	1.158	1.422

The longest predicted flight time was obtained when A = 1, B = 1, C = 0, and D = 0, corresponding to long rotors, long legs, narrow leg width, and no paper clip. This design had an estimated mean flight time of 1.965 seconds (95% CI: 1.833, 2.097). Therefore, this combination is our recommended configuration for maximizing helicopter flight duration.

### 3 Appendix

For Question 3, we first tried manually resetting reference levels. Results were similar and the conclusion were the same, but we ultimately decided to use the more efficient `emmeans`.

```
## Average flight time for factor combination (1) with factors (-, -, -, -):
## Estimate: 1.454 seconds
## 95% CI: ( 1.318 , 1.59 ) seconds

## Average flight time for factor combination a with factors (+, -, -, -):
## Estimate: 1.617 seconds
## 95% CI: ( 1.48 , 1.753 ) seconds

## Average flight time for factor combination b with factors (-, +, -, -):
## Estimate: 1.517 seconds
## 95% CI: ( 1.38 , 1.653 ) seconds

## Average flight time for factor combination c with factors (-, -, +, -):
## Estimate: 1.775 seconds
## 95% CI: ( 1.639 , 1.911 ) seconds

## Average flight time for factor combination d with factors (-, -, -, +):
## Estimate: 1.693 seconds
## 95% CI: ( 1.557 , 1.83 ) seconds

## Average flight time for factor combination ab with factors (+, +, -, -):
## Estimate: 1.994 seconds
## 95% CI: ( 1.858 , 2.13 ) seconds

## Average flight time for factor combination ac with factors (+, -, +, -):
## Estimate: 1.762 seconds
## 95% CI: ( 1.626 , 1.899 ) seconds

## Average flight time for factor combination ad with factors (+, -, -, +):
## Estimate: 1.869 seconds
## 95% CI: ( 1.733 , 2.005 ) seconds

## Average flight time for factor combination bc with factors (-, +, +, -):
## Estimate: 1.533 seconds
## 95% CI: ( 1.396 , 1.669 ) seconds

## Average flight time for factor combination bd with factors (-, +, -, +):
## Estimate: 1.359 seconds
## 95% CI: ( 1.223 , 1.495 ) seconds

## Average flight time for factor combination cd with factors (-, -, +, +):
## Estimate: 1.384 seconds
## 95% CI: ( 1.248 , 1.52 ) seconds

## Average flight time for factor combination abc with factors (+, +, +, -):
## Estimate: 1.825 seconds
## 95% CI: ( 1.689 , 1.961 ) seconds

## Average flight time for factor combination abd with factors (+, +, -, +):
## Estimate: 1.585 seconds
## 95% CI: ( 1.449 , 1.721 ) seconds

## Average flight time for factor combination acd with factors (+, -, +, +):
## Estimate: 1.585 seconds
## 95% CI: ( 1.449 , 1.721 ) seconds
```

```
## Average flight time for factor combination bcd with factors (-, +, +, +):  
## Estimate: 1.32 seconds  
## 95% CI: ( 1.184 , 1.456 ) seconds  
  
## Average flight time for factor combination abcd with factors (+, +, +, +):  
## Estimate: 1.561 seconds  
## 95% CI: ( 1.425 , 1.697 ) seconds
```