

Stir It Hot: Evaluating the Combined Impact of Temperature and Stirring on Salt Dissolution Time Using a Mixed-Effects Model

23070164

SUMMARY

This study explores how water temperature and stirring speed affect salt dissolution time. A 2×2 factorial design was employed to evaluate two treatment factors: HotterWater and FasterStir, each at two levels. We applied a Linear Mixed-effects Model (LMM) to quantify the main and interaction effects of the treatments while accounting for random variation between Operators. A total of 120 observations were collected in a completely randomized design with balanced replication. Operator variability was modelled as a random effect to enhance generalisability. Exploratory analysis revealed strong effects for both higher temperature and faster stirring, significantly reducing dissolution time. The interaction between these factors was statistically significant, suggesting that the combined effects are greater than additive. Residual analysis supported the assumptions of the ANOVA model, and power analysis confirmed excellent statistical strength. These findings indicate that temperature and stirring speed jointly influence dissolution time in a non-additive way. Post-hoc analysis using Tukey HSD further confirmed significant differences between all treatment combinations. Future studies could explore a broader set of conditions or include additional process variables.

1. INTRODUCTION

This study investigates how water temperature (HotterWater) and stirring speed (FasterStir) influence TotalSeconds, the time taken for 1 gram of salt to fully dissolve. A 2×2 factorial design was employed, with each factor tested at two levels: Low and High.

The experiment was conducted by three operators. However, as the aim was to assess the effects of HotterWater and FasterStir—not compare individual operators—Operator was treated as a random effect rather than fixed. This approach accounts for minor variation across experimenters and enhances the model's generalisability, aligning with established guidance that random effects are appropriate when factor levels reflect a broader population (Montgomery, 2017).

To appropriately partition variability, a Linear Mixed-Effects Model (LMM) was used. This separates:

1. Between-group variation from HotterWater and FasterStir
2. Within-group variation due to random operator differences

If Operator variance is small, it supports the conclusion that HotterWater and FasterStir are the primary contributors to changes in dissolution time, affirming the model's validity and broader relevance. Understanding dissolution rates has implications across fields like culinary science, pharmaceuticals, and chemical engineering, where solubility affects product quality. A factorial design enables efficient and simultaneous testing of multiple factors and their interaction, offering richer insight than varying one factor at a time.

2. EXPERIMENTAL DESIGN AND DATA COLLECTION

This experiment used a 2×2 factorial design, testing HotterWater and FasterStir at two levels each. Trials were conducted at room temperature (18°C) with consistent measurement protocols. Each operator conducted 10 repetitions per condition, yielding 120 observations total (30 per group).

2(a). Variables and Factors

1. HotterWater (Factor A): 0 = 18°C, 1 = 70°C
2. FasterStir (Factor B): 0 = 1 stir per 10s, 1 = 1 stir per 5s

3. TotalSeconds (Response): Time (in seconds) for 1 gram table salt (the type of salt used in this experiment) to dissolve
4. Operator (Random Effect): Three individuals, modeled as random to account for variability

2(b). Data Collection and Randomization Strategy

A completely randomized design was employed, ensuring:

1. Each operator followed a different random trial sequence
2. Balanced replication (10 trials per condition) ensured 30 observations per group
3. Measurements were standardized using the same tools for temperature, table salt, stirring, and timing

This structure minimized bias and upheld assumptions required for valid inference. A sample size of 30 per group was specifically chosen to fulfill ANOVA assumptions related to normality and variance homogeneity, enhancing statistical robustness (Field, 2013).

3. EXPLORATORY DATA ANALYSIS, MODEL FITTING, AND ASSUMPTION CHECKS

3(a). Exploratory Data Analysis (EDA) and Outlier Detection (Figure 1, Table S1, Table S2)

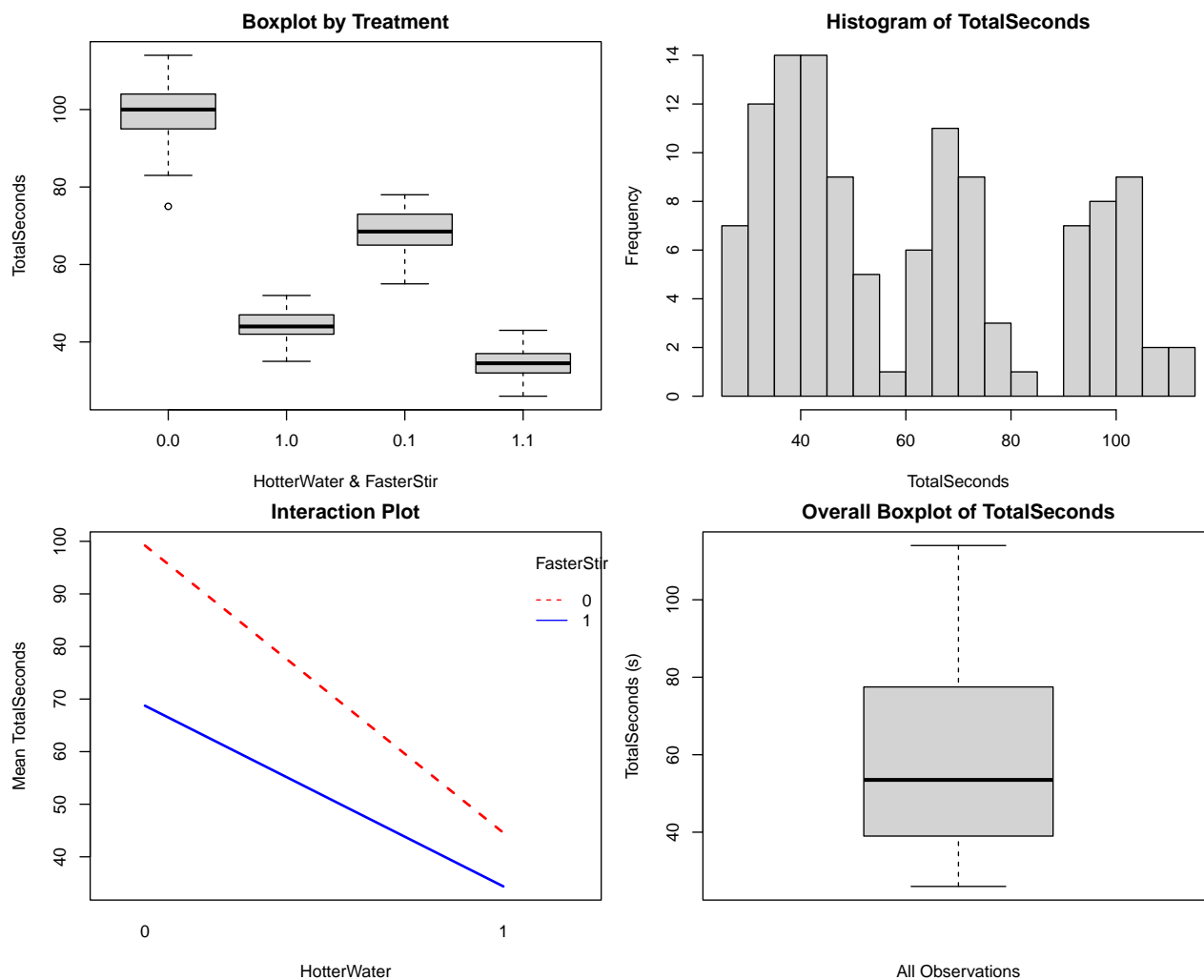


Figure 1: Exploratory Data Analysis (EDA) and Outlier Detection

Figure 1 displays four EDA plots. The histogram (top right) shows a slightly right-skewed distribution, with a mean of 61.7s and median of 53.5s—consistent with the summary in Table S1 (8. SUPPLEMENTARY INFORMATION). The treatment-wise boxplot (top left) reveals large group differences: the slowest group

(HotterWater = 0, FasterStir = 0) had a median near 99s, while the fastest (HotterWater = 1, FasterStir = 1) was around 34s, mirroring Table S2 group means.

The interaction plot (bottom left) displays non-parallel lines, indicating that the effect of one factor depends on the other. This visual supports the decision to include an interaction term in the model. The overall boxplot (bottom right) shows one mild outlier, but the spread aligns with the overall distribution in Table S1.

Together, these visuals and summary statistics confirm strong treatment effects and suggest interaction, justifying the model structure used in the analysis.

3(b). Normality, Homogeneity of Variance(ANOVA Assumption Checks Before Model Fitting) (Table S3, Table S4, Figure S1)

Shapiro-Wilk tests across the four treatment groups (Table S3) produced p-values comfortably above 0.05, suggesting normality within groups. This is supported by the Q-Q plots in Figure S1, which showed close alignment with the theoretical distribution. Levene's test for equal variances (Table S4), centred on the median, returned a p-value of 0.056—just above the 0.05 threshold. While this hints at some heterogeneity, it is not statistically significant.

Given the balanced group sizes, and the robustness of ANOVA-type models to minor variance inequality under such conditions (Field, 2013), we proceed assuming equal variances are reasonable.

3(c). Power Analysis (Table S5, Table S6)

Power analysis was conducted to confirm that the study design could detect meaningful effects. Two complementary approaches were used.

First, a prospective analysis determined the minimum sample size needed to achieve 90% power. Using standard deviation estimates from Idaho National Laboratory data (Breckenridge et al., 1998) ranging from 4–13 seconds, we conservatively chose $\sigma = 13$ and defined a meaningful difference of $\delta = 10$ seconds (from our experiment). This gave an effect size of $f = 0.77$. Under these assumptions, Table S5 shows that only 10 observations per group are required to reach 90% power at $\alpha = 0.05$.

Second, the actual power achieved under the current design (30 observations per group across four groups) was calculated. Table S6 confirms that the achieved power is extremely high (0.99995), indicating that the experiment is more than adequately powered. This two-pronged analysis supports the statistical validity of the findings.

4. LINEAR MIXED-EFFECTS MODEL ANALYSIS

4(a). Linear Mixed-Effects ANOVA Model

Given the variation introduced by different operators, a linear mixed-effects model (LMM) is applied as follows:

$$Y_{ijkl} = \mu + A_i + B_j + (AB)_{ij} + u_k + \epsilon_{ijkl}$$

Where:

- Y_{ijkl} : Observed TotalSeconds for the i th level of HotterWater and j th level of FasterStir, by operator k
- μ : Overall mean
- A_i : Fixed effect of HotterWater
- B_j : Fixed effect of FasterStir
- $(AB)_{ij}$: Interaction effect between HotterWater and FasterStir
- $u_k \sim N(0, \sigma_u^2)$: Random effect of Operator
- $\epsilon_{ijkl} \sim N(0, \sigma^2)$: Residual error

a. Define the Hypotheses

- $H_0^A : \mu_{A=0} = \mu_{A=1}, H_1^A : \mu_{A=0} \neq \mu_{A=1}$
- $H_0^B : \mu_{B=0} = \mu_{B=1}, H_1^B : \mu_{B=0} \neq \mu_{B=1}$
- $H_0^{AB} : (AB)_{ij} = 0, H_1^{AB} : (AB)_{ij} \neq 0$

b. Level of Significance

$$\alpha = 0.05$$

c. Decision Rule

Reject H_0 if the corresponding p-value < 0.05 .

d. Statistical Test and Output

A Type III ANOVA is performed on the fitted LMM. The F-statistics and p-values are used to assess the significance of each term.

Mixed-Effects Model Summary (Table S7)

The linear mixed-effects model showed that both HotterWater and FasterStir significantly reduced dissolution time. Raising temperature from 18°C to 70°C decreased time by 54.63 seconds, while faster stirring added a 30.47 second reduction. The interaction term (estimate = 20.27, $p < 0.001$) was also significant, indicating a synergistic effect beyond simple addition.

Operator was modelled as a random effect, with variance (1.06) much smaller than the residual (31.10), suggesting minimal inter-operator variation. This justifies its treatment as a random factor and improves generalisability beyond the three individuals tested, in line with recommendations from Field (2013) and Montgomery (2017).

Overall, the model supports the conclusion that both factors accelerate dissolution, and their combined effect is amplified. The small random variance and large, significant fixed effects affirm the model's robustness.

ANOVA Confirmation of Fixed Effects (Table 1)

Type III ANOVA results confirm all fixed effects as highly significant: HotterWater ($F = 1909.95$), FasterStir ($F = 398.77$), and their interaction ($F = 99.04$), all $p < 0.001$. The strong interaction reinforces its inclusion, as it shows that the impact of one factor depends on the other.

The use of Type III sums of squares is appropriate here, as it allows for accurate estimation and testing in a balanced design with interactions. The agreement between ANOVA and model output strengthens the overall conclusions.

Table 1. Type III ANOVA Test for Fixed Effects

Type III Analysis of Variance Table with Satterthwaite's method

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
HotterWater	59408	59408	1	114	1909.949	< 2.2e-16 ***
FasterStir	12403	12403	1	114	398.767	< 2.2e-16 ***
HotterWater:FasterStir	3081	3081	1	114	99.039	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

4(b). Residual Analysis (after fitting the model) (Table S8, Figure 2)

Residual diagnostics were conducted to assess model assumptions. The Shapiro-Wilk test (Table S8) yielded a p-value of 0.0005, indicating mild non-normality. However, with balanced groups ($n = 30$), the model remains robust, as supported by Field (2013) and Glass et al. (1972), who noted the F-test's reliability under moderate non-normality in balanced designs.

Diagnostic plots (Figure 2) showed no major violations: residuals appeared approximately normal, spread was consistent, and no funnel pattern or curvature was observed. Residuals also showed similar variance across treatment groups, supporting homoscedasticity.

Overall, no data points were removed, no transformation was needed, and the assumptions are sufficiently met for valid inference.

4(c).Post-Hoc Analysis: Pairwise Comparisons (Table 2)

Tukey HSD post-hoc analysis confirmed that all treatment combinations differed significantly ($p < 0.0001$;

Table 2). The largest contrast—between the high-temperature, fast-stirring group and the low-temperature, slow-stirring group—showed a 64.8 second reduction, demonstrating the strong combined effect.

Intermediate contrasts also revealed meaningful effects. A 54.6 second reduction was observed when increasing only temperature, and a 30.5 second reduction when increasing only stirring. Even between the two hot-water groups, faster stirring reduced dissolution time by 10.2 seconds.

These results align with the mixed model estimates and reinforce the strong main and interaction effects. The Tukey-adjusted comparisons highlight the factorial structure of the experiment and complement the overall ANOVA findings.

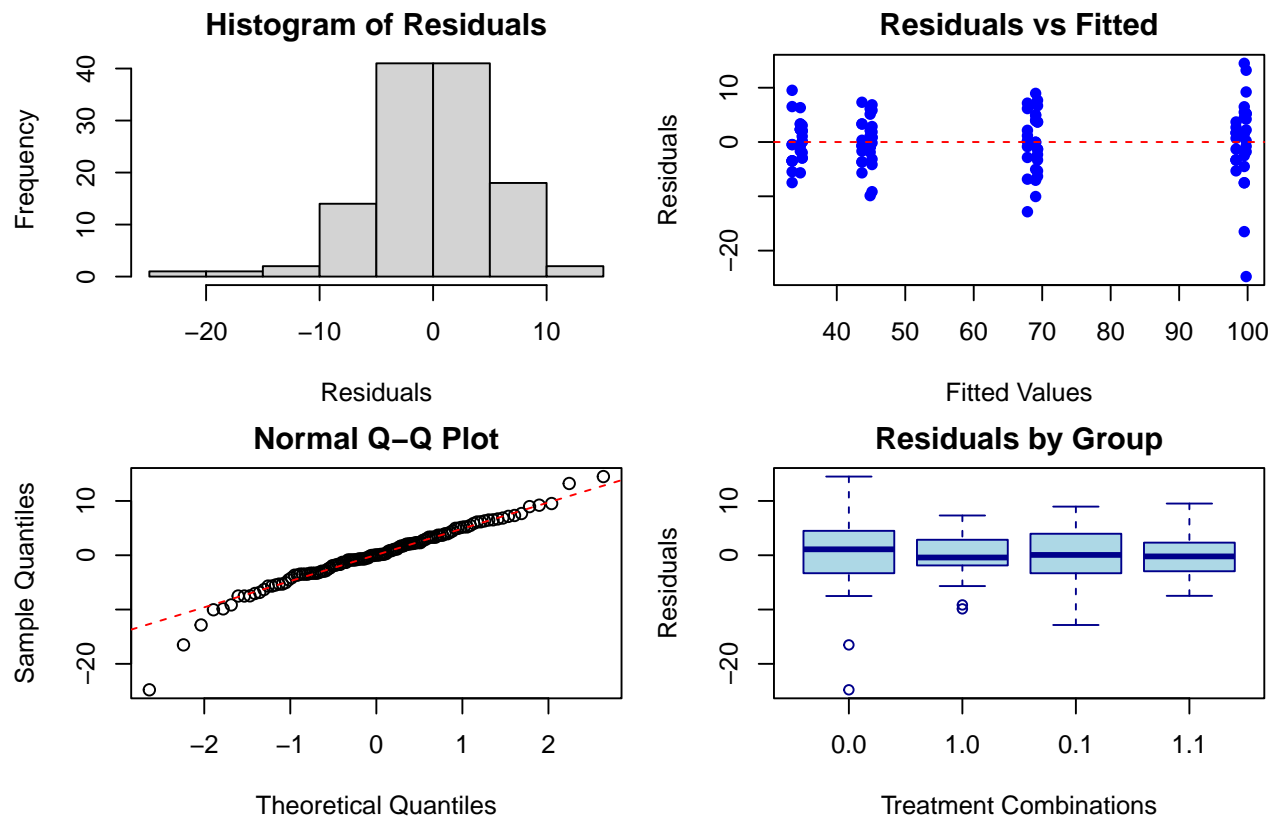


Figure 2: Residual Diagnostics/Check

****Table 2. Tukey HSD Test Result (Adjusted Pairwise Comparisons for Interaction)****

contrast	estimate	SE	df	t.ratio	p.value
HotterWater0 FasterStir0 - HotterWater1 FasterStir0	54.6	1.44	114	37.940	<.0001
HotterWater0 FasterStir0 - HotterWater0 FasterStir1	30.5	1.44	114	21.157	<.0001
HotterWater0 FasterStir0 - HotterWater1 FasterStir1	64.8	1.44	114	45.023	<.0001
HotterWater1 FasterStir0 - HotterWater0 FasterStir1	-24.2	1.44	114	-16.782	<.0001
HotterWater1 FasterStir0 - HotterWater1 FasterStir1	10.2	1.44	114	7.083	<.0001
HotterWater0 FasterStir1 - HotterWater1 FasterStir1	34.4	1.44	114	23.866	<.0001

Degrees-of-freedom method: kenward-roger

P value adjustment: tukey method for comparing a family of 4 estimates

5. DISCUSSION & CONCLUSION

The findings clearly support the study's aims, showing strong and significant main and interaction effects of temperature and stirring on salt dissolution. These results are consistent across model coefficients, ANOVA outputs, and post-hoc comparisons. The Tukey HSD test (Table 2) confirmed that all treatment contrasts were significant ($p < 0.0001$), reinforcing these effects.

Crucially, the small variance attributed to Operator (Table S7) supports modelling it as a random effect. This suggests that HotterWater and FasterStir are the primary drivers of dissolution time, making the model both accurate and generalisable. It also aligns with the principles of factorial design, which aim to isolate treatment effects while accounting for minimal nuisance variability.

These findings have practical relevance for improving dissolution processes in fields like food production, pharmaceuticals, and manufacturing—where optimising temperature and stirring can boost efficiency and reduce cost. However, testing only two levels per factor limits conclusions about broader or nonlinear effects. Standardised lab conditions may also miss variability present in real-world settings.

Future studies could explore finer or continuous levels and consider real-world constraints such as implementation practicality and equipment compatibility. Overall, the evidence shows that hotter water and faster stirring both independently and jointly accelerate salt dissolution. The conclusions are robust, though additional research is needed to confirm their generalisability.

6. REFERENCES

1. Breckenridge, R. P., Arnold, B. W., & Kohout, R. (1998). Solubility and Dissolution Rate of LiCl-KCl-NaCl. Idaho National Laboratory. Retrieved from https://inldigitallibrary.inl.gov/sites/sti/sti/Sort_117419.pdf

2. Department of Statistical Science, University College London. (2024). *STAT0029: Experimental Design and Analysis – Lecture Notes*. UCL Moodle.

3. Field, A. (2013). *Discovering Statistics Using R*. SAGE Publications.

4. Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, 42(3), 237–288.

5. Montgomery, D.C. (2017). *Design and Analysis of Experiments*. 9th Edition. Wiley.

7. WORD COUNT

(1757)

8. SUPPLEMENTARY INFORMATION

****Table S1. Summary Statistics of TotalSeconds****

	Statistic	Value
1	Min.	26.00000
2	1st Qu.	39.50000
3	Median	53.50000
4	Mean	61.71667
5	3rd Qu.	77.25000
6	Max.	114.00000

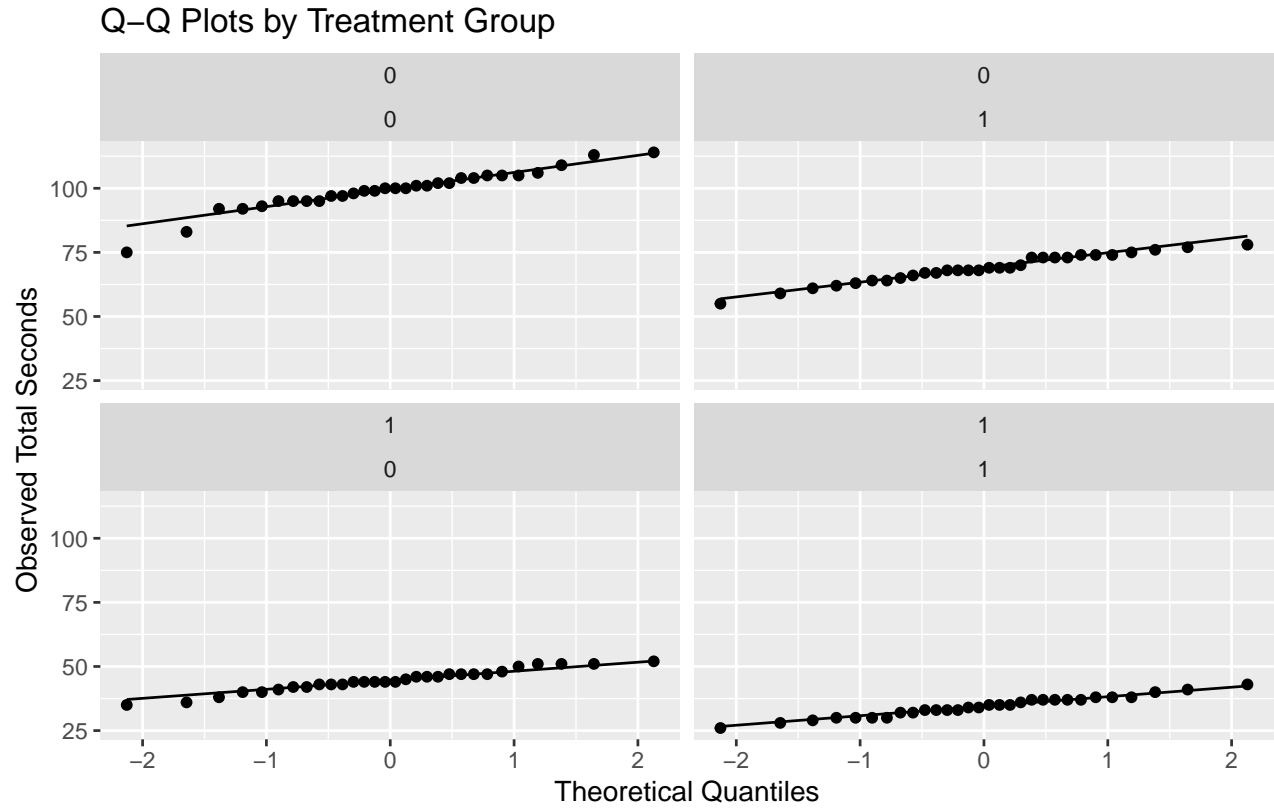
****Table S2. Group Means by Treatment****

	HotterWater	FasterStir	TotalSeconds
1	0	0	99.20000
2	1	0	44.56667
3	0	1	68.73333
4	1	1	34.36667

****Table S3. Shapiro-Wilk Test per Group****

```
# A tibble: 4 x 4
  HotterWater FasterStir      W p_value
  <fct>      <fct>      <dbl> <dbl>
1 0          0          0.936 0.0699
2 0          1          0.970 0.541
3 1          0          0.969 0.515
4 1          1          0.983 0.904
```

****Figure S1. Q-Q Plots by Treatment Group****



****Table S4. Levene’s Test for Equal Variance****

```
Levene’s Test for Homogeneity of Variance (center = median)
      Df F value  Pr(>F)
group   3  2.5878 0.05637 .
    116
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

****Table S5. Minimal Sample Size per Group to Achieve 90% Power****

```
Balanced one-way analysis of variance power calculation

      k = 2
      n = 9.945319
      f = 0.7692308
sig.level = 0.05
power = 0.9
```

NOTE: n is number in each group

****Table S6. Power Calculation for Current Design (n = 30 per group)****

Balanced one-way analysis of variance power calculation

```

k = 2
n = 30
f = 0.7692308
sig.level = 0.05
power = 0.9999514

```

NOTE: n is number in each group

****Table S7. Linear Mixed-Effects Model Summary (Full Output)****

Linear mixed model fit by REML. t-tests use Satterthwaite's method [`'lmerModLmerTest'`]

Formula: TotalSeconds ~ HotterWater * FasterStir + (1 | Operator)

Data: data_experiment

REML criterion at convergence: 743.3

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.4435	-0.5718	0.0092	0.5963	2.5985

Random effects:

Groups	Name	Variance	Std.Dev.
Operator	(Intercept)	1.061	1.030
Residual		31.104	5.577

Number of obs: 120, groups: Operator, 3

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	99.200	1.179	10.010	84.122	1.34e-15 ***
HotterWater1	-54.633	1.440	114.000	-37.940	< 2e-16 ***
FasterStir1	-30.467	1.440	114.000	-21.157	< 2e-16 ***
HotterWater1:FasterStir1	20.267	2.036	114.000	9.952	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	HttrW1	FstrS1
HotterWat1	-0.611		
FasterStir1	-0.611	0.500	
HttrWt1:FS1	0.432	-0.707	-0.707

****Table S8. Shapiro Test for Normality of Residual****

Shapiro-Wilk normality test

```

data: resid(fit_mixed)
W = 0.95512, p-value = 0.0005231

```


DATASET

The dataset used in this analysis is provided as a separate file named **experiment_data.csv**. It should be placed in the working directory before running the R code. Below are the details of the dataset:

TotalSeconds (Numeric): Response variable measuring time in seconds.

HotterWater (Factor, Levels: “Low”, “High”): Categorical variable representing water temperature.

FasterStir (Factor, Levels: “Low”, “High”): Categorical variable representing stirring speed.

Operator (Factor): Represents the individual who conducted the experiment.

The Table below presents the raw data used in the analysis. It includes the Operator ID, treatment levels (HotterWater and FasterStir), and the response variable (TotalSeconds). The data are displayed in two columns to improve readability.

Notes on Variable Names in CSV

- **Operator:** The person conducting the trial (coded as 1, 2, 3)
- **HotterWater:** Temperature treatment (0 = 18°C, 1 = 70°C)
- **FasterStir:** Stirring frequency (0 = every 10 seconds, 1 = every 5 seconds)
- **TotalSeconds:** Time (in seconds) for the salt to fully dissolve

Table 1: Raw experimental data. HW = HotterWater, FS = FasterStir. Op=Operator

Row	Op	HW	FS	Time	Row2	Op2	HW2	FS2	Time2
1	1	1	1	37	61	1	1	0	52
2	2	1	1	40	62	2	1	0	47
3	3	1	1	37	63	3	1	0	50
4	1	1	0	42	64	1	1	1	35
5	2	1	0	42	65	2	1	1	28
6	3	1	0	35	66	3	1	1	37
7	1	1	0	46	67	1	1	0	47
8	2	1	0	47	68	2	1	0	43
9	3	1	0	43	69	3	1	0	47
10	1	1	0	41	70	1	1	1	38
11	2	1	0	43	71	2	1	1	33
12	3	1	1	41	72	3	1	1	37
13	1	1	1	35	73	1	1	0	45
14	2	1	1	33	74	2	1	0	40
15	3	1	0	44	75	3	1	0	44
16	1	0	1	66	76	1	1	1	32
17	2	0	1	74	77	2	1	1	30
18	3	0	1	69	78	3	1	1	33
19	1	0	0	75	79	1	0	1	68
20	2	0	0	93	80	2	0	1	70
21	3	0	0	92	81	3	0	1	78
22	1	0	1	76	82	1	0	0	104
23	2	0	1	69	83	2	0	0	95
24	3	0	1	69	84	3	0	0	106
25	1	0	0	99	85	1	0	1	73
26	2	0	0	102	86	2	0	1	68
27	3	0	0	114	87	3	0	1	74
28	1	0	1	67	88	1	0	0	105
29	2	0	1	55	89	2	0	0	100
30	3	0	1	73	90	3	0	0	105
31	1	0	0	102	91	1	0	1	77

Row	Op	HW	FS	Time	Row2	Op2	HW2	FS2	Time2
32	2	0	0	100	92	2	0	1	68
33	3	0	0	104	93	3	0	1	73
34	1	1	1	37	94	1	0	0	100
35	2	1	1	43	95	2	0	0	95
36	3	1	1	38	96	3	0	0	97
37	1	1	0	51	97	1	0	1	68
38	2	1	0	51	98	2	0	1	75
39	3	1	0	51	99	3	0	1	74
40	1	1	1	35	100	1	0	0	113
41	2	1	1	30	101	2	0	0	95
42	3	1	1	38	102	3	0	0	95
43	1	1	0	36	103	1	0	1	73
44	2	1	0	44	104	2	0	1	67
45	3	1	0	46	105	3	0	1	62
46	1	1	1	32	106	1	0	0	105
47	2	1	1	26	107	2	0	0	99
48	3	1	1	29	108	3	0	0	92
49	1	1	0	46	109	1	0	1	64
50	2	1	0	38	110	2	0	1	65
51	3	1	0	44	111	3	0	1	59
52	1	1	1	33	112	1	0	0	109
53	2	1	1	30	113	2	0	0	97
54	3	1	1	34	114	3	0	0	83
55	1	1	0	48	115	1	0	1	63
56	2	1	0	40	116	2	0	1	61
57	3	1	0	44	117	3	0	1	64
58	1	1	1	36	118	1	0	0	98
59	2	1	1	30	119	2	0	0	101
60	3	1	1	34	120	3	0	0	101

The dataset includes 120 observations across 4 experimental conditions. Displayed above in 2 parts to preserve formatting, each row includes operator ID, treatment levels, and response time.

CODE

```
knitr::opts_chunk$set(
  echo = TRUE,
  fig.cap = TRUE,
  fig.width = 7,
  comment = "",
  tidy = FALSE,
  warning = FALSE,
  message = FALSE
)
options(width = 90)
suppressWarnings(suppressMessages({
  library(ggplot2)
  library(dplyr)
  library(gridExtra)
  library(lme4)
  library(car)
  library(pwr)
  library(lmtest)
  library(lmerTest)
```

```

library(emmeans)
library(knitr)

}))

set.seed(123) # set seed

#Read the experiment data
data_experiment <- read.csv("experiment_data.csv")
data_experiment$Operator <- as.factor(data_experiment$Operator)
data_experiment$HotterWater <- as.factor(data_experiment$HotterWater)
data_experiment$FasterStir <- as.factor(data_experiment$FasterStir)

## ==1. EDA *(Supplementary Output Table S1 and S2)*==

#Make summary statistics
summary_tbl <- data.frame(Statistic = names(summary
                                     (data_experiment$TotalSeconds)),
                          Value = as.numeric
                               (summary(data_experiment$TotalSeconds)))

group_means <- aggregate(TotalSeconds ~ HotterWater + FasterStir,
                          data = data_experiment, mean)

#Make 4in1 plot
par(mfrow = c(2, 2))
par(mar=c(4,4,2,1))

# 1a. Boxplot by treatment
boxplot(data_experiment$TotalSeconds ~ interaction(data_experiment$HotterWater,
                                                    data_experiment$FasterStir),
        main = "Boxplot by Treatment", xlab = "HotterWater & FasterStir",
        ylab = "TotalSeconds")

# 1b. Histogram
hist(data_experiment$TotalSeconds, breaks = 20, col = "lightgray",
     main = "Histogram of TotalSeconds", xlab = "TotalSeconds")

# 1c. Interaction plot
interaction.plot(data_experiment$HotterWater, data_experiment$FasterStir,
                 data_experiment$TotalSeconds,
                 col = c("red", "blue"), lwd = 2,
                 xlab = "HotterWater", ylab = "Mean TotalSeconds",
                 main = "Interaction Plot", trace.label = "FasterStir")

# 1d. Overall boxplot
boxplot(data_experiment$TotalSeconds,
        main = "Overall Boxplot of TotalSeconds",
        ylab = "TotalSeconds (s)",
        xlab = "All Observations")

## ==2. ASSUMPTION CHECKS *(Supplementary Output Table S3)*==

#2a. Shapiro and Q-Q Plot for normality tests
grouped_shapiro <- data_experiment %>%
  group_by(HotterWater, FasterStir) %>%

```

```

summarise(
  W = shapiro.test(TotalSeconds)$statistic,
  p_value = shapiro.test(TotalSeconds)$p.value,
  .groups = "drop"
)

#2b. Q-Q Plots by Group-Figure S1
qq_plot_normal <- ggplot(data_experiment, aes(sample = TotalSeconds)) +
  stat_qq() + stat_qq_line() +
  facet_wrap(~ HotterWater + FasterStir) +
  labs(
    title = "Q-Q Plots by Treatment Group",
    x = "Theoretical Quantiles",
    y = "Observed Total Seconds"
  )

#2b. Levene Test
levene_result <- leveneTest(TotalSeconds ~ HotterWater * FasterStir,
  data = data_experiment)

# ===3. Power Analysis (Required & Achieved) Table===
sigma <- 13          # Standard deviation assumption from prior study
delta <- 10          # Smallest effect size from this experiment
alpha <- 0.05        # Significance level
k <- 2              # Number of factors
f <- delta / sigma   # Cohen's f
desired_power <- 0.9 # Desired power

# 3a. Calculate required sample size
power_needed <- pwr.anova.test(k = k, f = f, sig.level = alpha,
  power = desired_power)

# 3b. Calculate achieved power for current design (n = 30 per group)
power_check <- pwr.anova.test(k = k, n = 30, f = f, sig.level = alpha)

# ===4. Fit the linear mixed-effects model===
fit_mixed <- lmer(TotalSeconds ~ HotterWater * FasterStir + (1 | Operator),
  data = data_experiment)

anova_result <- anova(fit_mixed, type = 3)

# ===5. Residual Check===

opar <- par(mfrow = c(2, 2), mar = c(4, 4, 2, 1))
# Histogram of residuals
hist(resid(fit_mixed),
  main = "Histogram of Residuals",
  xlab = "Residuals",
  col = "lightgray", border = "black")

# Residuals vs Fitted values
plot(fitted(fit_mixed), resid(fit_mixed),
  main = "Residuals vs Fitted",
  xlab = "Fitted Values",
  ylab = "Residuals",

```

```

col = "blue", pch = 16)
abline(h = 0, col = "red", lty = 2)

# Q-Q plot
qqnorm(resid(fit_mixed), main = "Normal Q-Q Plot")
qqline(resid(fit_mixed), col = "red", lty = 2)

# Boxplot of residuals by treatment combinations(optional but seen in exercises)
boxplot(resid(fit_mixed) ~ interaction(data_experiment$HotterWater,
                                     data_experiment$FasterStir),
        xlab = "Treatment Combinations", ylab = "Residuals",
        main = "Residuals by Group",
        col = "lightblue", border = "darkblue")

# ===6. Estimated marginal means for all combinations===
emm <- emmeans(fit_mixed, ~ HotterWater * FasterStir)
tukey_result <- pairs(emm, adjust = "tukey")

#===Print All Output===

cat("**Table S1. Summary Statistics of TotalSeconds**\n\n")
print(summary_tbl)

cat("**Table S2. Group Means by Treatment**\n\n")
print(group_means )

cat("**Table S3. Shapiro-Wilk Test per Group**\n\n")
print(grouped_shapiro )

cat("**Figure S1. Q-Q Plots by Group**\n\n")
print(qq_plot_normal)

#Levene's Test for Equal Variance Table S4
cat("**Table S4. Levene's Test for Equal Variance**\n\n")
print(levene_result)

#Table S5: Print result for required sample size
cat("**Table S5. Minimal Sample Size per Group to Achieve 90% Power**\n\n")
print(power_needed)

#Table S6. Power Calculation for Current Design
cat("\n**Table S6. Power Calculation for Current Design
    (n = 30 per group)**\n\n")
print(power_check)

#Table S7. Linear Mixed-Effects Model Summary (Full Output)
cat("**Table S7. Linear Mixed-Effects Model Summary (Full Output)**\n\n")
print(summary(fit_mixed))

#Table S8. Shapiro Test for Normality of Residual
cat("\n**Table S8. Shapiro Test for Normality of Residual**\n\n")
shapiro.test(resid(fit_mixed))

#Table 1. Type III ANOVA Test for Fixed Effects

```

```
cat("\n**Table 1. Type III ANOVA Test for Fixed Effects**\n\n")
print(anova_result)

#Table 2. Tukey HSD Test Result (Adjusted Pairwise Comparisons for Interaction)
cat("**Table 2. Tukey HSD Test Result (Adjusted Pairwise Comparisons
    for Interaction)**\n\n")
print(summary(tukey_result))
```