**Experiment Project**

**April 12th, 2020**

**Connor Armstrong**

**STAT 8120 – Applied Experimental Design – Dr. Victor Kane**

**EGP Count = 8, Footnotes in exponentiated brackets [ ]**

*Abstract: An experiment was developed to determine how and if a series of factors affect response time for a command to be obeyed by dogs. The purpose of the experiment is to fulfill the experimental project requirement for STAT 8120 – Applied Experimental Design, by successfully implementing Experimental Good Practices (EGP) learned throughout the Spring 2020 semester. A number of factors were determined to be significant in reducing the response time of dogs to a command.*

**Section 1: Background**

The behavior of dogs has been studied in a range of manners. The Animal Behaviour Cognition & Welfare Group from the University of London published a paper, *Factors affecting response of dogs to obedience instruction a field and experimental study*, wherein it was determined that dogs have more difficulties ‘generalizing’ novel commands versus well established ones (Reference 6.1). The Center for Shelter Dogs in the Department of Clinical Sciences from the Cummings School of Veterinary Medicine at Tufts University published a study, *Characteristics of Excitable Dog Behavior Based on Owners’ Report from a Self-Selected Study,* focused on the excitability of shelter dogs and how that impacts the probability that the dog is kept (Reference 6.2).

When performing experiments on animals, special care must be taken to ensure that the experimental conditions are consistent and repeatable. Controlling for nuisance factors will be instrumental to the viability and accuracy of conclusions.

**Section 2: Experiment Design**

**2.1 Experimental Factors**

1. Factors of Interest[1]
   1. Treats – treat in hand/not in hand
   2. Location – inside/outside
   3. Operator – Operator 1, Operator 2, Operator 3
2. Nuisance Factors[5], Blocking[4] and run order independence[5]
   1. 2 Dogs – Dog 1 and Dog 2
   2. Day of Week – Friday, Saturday, Sunday
   3. Run Order[5] – Assume independence

**2.2 Response**

* Time to obey command to be **minimized** – “Sit”
  + Measured in seconds
  + Multiple commands allowed – must be consistent and at regular intervals
  + If dog does not obey command, reset run and do not record observation

**2.3 Additional Notes**

1. Standard order for factors was Randomized[2]
2. The following replicated 23 design was assigned to each of the 3 operators.
3. 2 replicates[3] \* 23 observations for 3 factors \* 3 operators = 48 total observations

**Table 2.3.1: Experimental Design Matrix**

|  |  |  |  |
| --- | --- | --- | --- |
| **Standard Order** | **Dog** | **Treat** | **Location** |
| 1 | 1 | 1 | 1 |
| 2 | 2 | 1 | 1 |
| 3 | 1 | 2 | 1 |
| 4 | 2 | 2 | 1 |
| 5 | 1 | 1 | 2 |
| 6 | 2 | 1 | 2 |
| 7 | 1 | 2 | 2 |
| 8 | 2 | 2 | 2 |
| 9 | 1 | 1 | 1 |
| 10 | 2 | 1 | 1 |
| 11 | 1 | 2 | 1 |
| 12 | 2 | 2 | 1 |
| 13 | 1 | 1 | 2 |
| 14 | 2 | 1 | 2 |
| 15 | 1 | 2 | 2 |
| 16 | 2 | 2 | 2 |

**Coded Factor Levels**

* Dog
  + Dog 1: (1)
  + Dog 2: (2)
* Treats
  + No Treat: (1)
  + Treat in Hand: (2)
* Location
  + Inside: (1)
  + Outside: (2)

**2.4 Hypothesis Testing**

The following hypotheses will be evaluated.

1. Holding a treat affects the mean time it takes for both dogs to obey the command.

H0: μtreat = μwithout; HA: μtreat ≠ μwithout

1. The location which the experiment is run influences the response.

H0: μinside = μoutside; HA: μinside ≠ μoutside

1. The ‘operator’ of the experiment influences the response.

H0: μoperator1 = μoperator2 = μoperator3; HA: The mean time to sit is not equivalent between the 3 operators

**Section 3: Measurement Evaluation**

**3.1 Response Measurement**

The response to be measured is the time between the command, “Sit”, to be fully and correctly executed. The command will be used after its attention has been gotten, using auditory signals such as “clicks” and referring to the dog by name, and eye contact made with the operator for each trial run. The timer will be started simultaneously with the command, and stopped upon full execution of the desired behavior. If the dog sits before the command is made, or if the dog appears to have anticipated the command, the trial will be reset before continuing with the experiment.

**3.2 Trial Experiment with One Set of Factor Conditions**

Having time restrictions to complete the experiment, the first 12 trials completed in the first day of experimentation will be analyzed as a modified trial experiment. Two sample t-tests will be analyzed for factors dog and treat. The full analysis will control for the factors of operator, day, location, and run order as well as second order interaction terms1.

Initial analysis of the t-tests for factors dog and treat indicate that the factor Dog appears significant while treat does not. More observations will improve the power and reliability of conclusions.

**Minitab t-Test Results for Dog 1 vs Dog 2**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Null hypothesis | | | H₀: μ₁ - µ₂ = 0 | |
| Alternative hypothesis | | | H₁: μ₁ - µ₂ ≠ 0 | |
| **T-Value** | **DF** | **P-Value** | |
| -2.62 | 8 | 0.031 | |

**Descriptive Statistics: Response (seconds)**

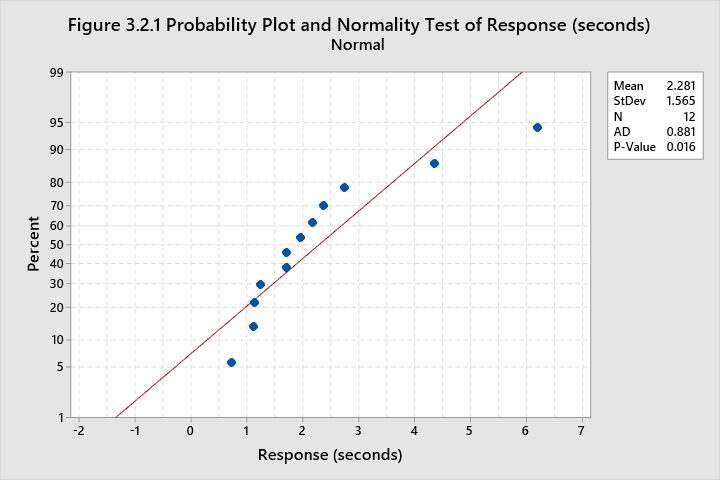
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Treat** | **N** | **Mean** | **StDev** | **SE Mean** |
| 1 | 4 | 1.195 | 0.403 | 0.20 |
| 2 | 8 | 2.82 | 1.66 | 0.59 |

**Minitab t-Test Results for Treat (1) vs No Treat (2)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Null hypothesis | | | H₀: μ₁ - µ₂ = 0 | |
| Alternative hypothesis | | | H₁: μ₁ - µ₂ ≠ 0 | |
| **T-Value** | **DF** | **P-Value** | |
| -1.16 | 5 | 0.299 | |

**Descriptive Statistics: Response (seconds)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Location** | **N** | **Mean** | **StDev** | **SE Mean** |
| 1 | 6 | 1.765 | 0.574 | 0.23 |
| 2 | 6 | 2.80 | 2.10 | 0.86 |



T-tests require validation of certain assumptions. The normal probability plot and Anderson-Darling normality test indicate that the response is not normal, which may indicate a need for transformation of the response variable, response time. Levene’s test of homogeneity of variance concludes that the variances between the two dogs are likely equivalent while the variances between treat and no treat are similarly significant at the standard α = 0.05 significance level.

**Levene’s Test for Dog**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Null hypothesis | | | H₀: σ₁ / σ₂ = 1 | | |
| Alternative hypothesis | | | H₁: σ₁ / σ₂ ≠ 1 | | |
| Significance level | | | α = 0.05 | | |
| **Method** | **Test Statistic** | **DF1** | | **DF2** | **P-Value** | |
| Levene | 1.48 | 1 | | 10 | 0.252 | |

**Levene’s Test for Treat**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Null hypothesis | | | H₀: σ₁ / σ₂ = 1 | | |
| Alternative hypothesis | | | H₁: σ₁ / σ₂ ≠ 1 | | |
| Significance level | | | α = 0.05 | | |
| **Method** | **Test Statistic** | **DF1** | | **DF2** | **P-Value** | |
| Levene | 1.44 | 1 | | 10 | 0.258 | |

**Section 4: Data Collection Plan[5]**

The full experimental design matrix containing all 48 observations can be seen in table 5.1. Each “Operator” was assigned a 16 observation experimental design matrix and the full set was randomized using Minitab. Data collection was performed by the operator selected for the observation to ensure that timing was as accurate as possible. The timer used for each observation was the same, using the default clock application on a Samsung Galaxy S20 5G smartphone. This timer records seconds to 2 decimal points. Operator reaction time will be assumed to be consistent with *A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students* from the Department of Physiology, Government Medical College, Patiala, Punjab, India. This study concluded that reaction times of some humans is approximately 200 milliseconds.

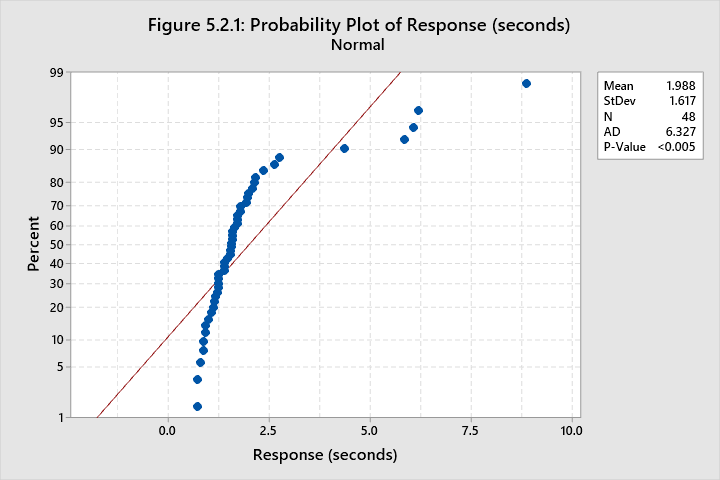
**Section 5: Results**

**5.1 Full Randomized Experimental Design Matrix with Results**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Run Order** | **Standard Order** | **Dog** | **Treat** | **Location** | **Operator** | **Response (seconds)** | **Date (mm/dd)** |
| 1 | 26 | 2 | 1 | 1 | 2 | 1.70 | 04/10 |
| 2 | 24 | 2 | 2 | 2 | 2 | 6.20 | 04/10 |
| 3 | 35 | 1 | 2 | 1 | 3 | 2.36 | 04/10 |
| 4 | 46 | 2 | 1 | 2 | 3 | 1.23 | 04/10 |
| 5 | 8 | 2 | 2 | 2 | 1 | 4.36 | 04/10 |
| 6 | 7 | 1 | 2 | 2 | 1 | 1.11 | 04/10 |
| 7 | 19 | 1 | 2 | 1 | 2 | 2.16 | 04/10 |
| 8 | 12 | 2 | 2 | 1 | 1 | 1.70 | 04/10 |
| 9 | 27 | 1 | 2 | 1 | 2 | 1.95 | 04/10 |
| 10 | 5 | 1 | 1 | 2 | 1 | 1.13 | 04/10 |
| 11 | 32 | 2 | 2 | 2 | 2 | 2.75 | 04/10 |
| 12 | 2 | 2 | 1 | 1 | 1 | 0.72 | 04/10 |
| 13 | 43 | 1 | 2 | 1 | 3 | 1.79 | 04/11 |
| 14 | 30 | 2 | 1 | 2 | 2 | 0.92 | 04/11 |
| 15 | 40 | 2 | 2 | 2 | 3 | 1.98 | 04/11 |
| 16 | 9 | 1 | 1 | 1 | 1 | 1.77 | 04/11 |
| 17 | 18 | 2 | 1 | 1 | 2 | 1.16 | 04/11 |
| 18 | 33 | 1 | 1 | 1 | 3 | 1.53 | 04/11 |
| 19 | 11 | 1 | 2 | 1 | 1 | 1.63 | 04/11 |
| 20 | 22 | 2 | 1 | 2 | 2 | 1.21 | 04/11 |
| 21 | 31 | 1 | 2 | 2 | 2 | 1.70 | 04/11 |
| 22 | 20 | 2 | 2 | 1 | 2 | 2.12 | 04/11 |
| 23 | 3 | 1 | 2 | 1 | 1 | 1.24 | 04/11 |
| 24 | 47 | 1 | 2 | 2 | 3 | 2.62 | 04/11 |
| 25 | 37 | 1 | 1 | 2 | 3 | 1.56 | 04/11 |
| 26 | 38 | 2 | 1 | 2 | 3 | 0.78 | 04/11 |
| 27 | 45 | 1 | 1 | 2 | 3 | 1.37 | 04/11 |
| 28 | 14 | 2 | 1 | 2 | 1 | 0.71 | 04/11 |
| 29 | 23 | 1 | 2 | 2 | 2 | 1.55 | 04/11 |
| 30 | 29 | 1 | 1 | 2 | 2 | 0.86 | 04/11 |
| 31 | 21 | 1 | 1 | 2 | 2 | 1.07 | 04/11 |
| 32 | 1 | 1 | 1 | 1 | 1 | 1.24 | 04/11 |
| 33 | 16 | 2 | 2 | 2 | 1 | 6.06 | 04/11 |
| 34 | 36 | 2 | 2 | 1 | 3 | 1.24 | 04/11 |
| 35 | 28 | 2 | 2 | 1 | 2 | 5.84 | 04/11 |
| 36 | 6 | 2 | 1 | 2 | 1 | 1.57 | 04/12 |
| 37 | 44 | 2 | 2 | 1 | 3 | 8.87 | 04/12 |
| 38 | 13 | 1 | 1 | 2 | 1 | 1.45 | 04/12 |
| 39 | 48 | 2 | 2 | 2 | 3 | 1.92 | 04/12 |
| 40 | 42 | 2 | 1 | 1 | 3 | 1.38 | 04/12 |
| 41 | 34 | 2 | 1 | 1 | 3 | 0.85 | 04/12 |
| 42 | 39 | 1 | 2 | 2 | 3 | 1.52 | 04/12 |
| 43 | 15 | 1 | 2 | 2 | 1 | 1.38 | 04/12 |
| 44 | 25 | 1 | 1 | 1 | 2 | 1.58 | 04/12 |
| 45 | 41 | 1 | 1 | 1 | 3 | 0.99 | 04/12 |
| 46 | 10 | 2 | 1 | 1 | 1 | 0.92 | 04/12 |
| 47 | 17 | 1 | 1 | 1 | 2 | 2.08 | 04/12 |
| 48 | 4 | 2 | 2 | 1 | 1 | 1.59 | 04/12 |

**5.2 Normalizing Response**

The response variable, response time, was checked for normality using the Anderson-Darling normality test. Having a p-value of less than 0.005, the response is not normal.



The Box-Cox transformation function in Minitab estimates the best transformation to normalize the response variable. In this case, the best transformation is to take the inverse of the response, ie y’ = 1/y.

**Box-Cox Transformation Method**

|  |  |
| --- | --- |
| Rounded λ | -1 |
| Estimated λ | -0.773582 |
| 95% CI for λ | (-1.14108, -0.431082) |

**5.3 Model Optimization**

The full ANOVA model with second-order interactions was computed using Minitab.

* The initial R2adj for the full model was determined to be 57.36%.
* Variance inflation factors for each factor and interaction lie below 10, which means that there is no significant issue with multicollinearity.
* There are 5 potential outliers in the full second-order model with the transformed response.
* Some factor interactions have very high p-values (above 0.8). Eliminating insignificant factor interactions may improve the accuracy and reliability of the model.

**Table 5.3.1: Model Optimization Process**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model Transformation** | **Removed Factors?** | **R2adj** | **A.D. Normality Test p-value** | **# of Potential Outliers** | **VIF>10?** |
| No Transformation | No | 25.45% | 0.022 | 3 | No |
| Inverse Transformation | No | 57.36% | 0.530 | 5 | No |
| Inverse Transformation | Treat\*Day, Treat\*Operator, Operator\*Day | 62.24% | 0.782 | 2 | No |
| Inverse Transformation | Same as above plus Treat\*Location | 63.40% | 0.785 | 2 | No |
| Inverse Transformation | Same as above plus Location\*Day | 63.63% | 0.691 | 3 | No |

**5.4 ANOVA for Selected Model**

The Model from table 5.3.1 highlighted in green will be selected for further analysis. This model will be used because it explains approximately as much as the model below it, but contains fewer potential outliers. Model selection was optimized to reduce ANOVA assumption violations. The ANOVA table for the selected model is below. Considering the standard significance level of α = 0.05, those factors and factor combinations which meet the criteria for significance are highlighted in table 5.4.1.

**Table 5.4.1 Analysis of Variance for Selected Model**

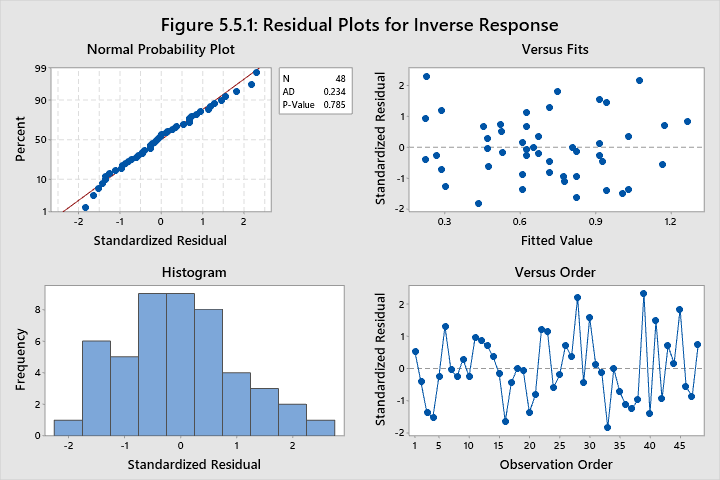
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Adj SS** | **Adj MS** | **F-Value** | **P-Value** |
| Dog | 1 | 0.04882 | 0.04882 | 1.41 | 0.244 |
| Treat | 1 | 1.73258 | 1.73258 | 50.02 | 0.000 |
| Location | 1 | 0.01442 | 0.01442 | 0.42 | 0.524 |
| Operator | 2 | 0.22993 | 0.11497 | 3.32 | 0.050 |
| Day | 2 | 0.07391 | 0.03696 | 1.07 | 0.357 |
| Dog\*Treat | 1 | 0.50051 | 0.50051 | 14.45 | 0.001 |
| Dog\*Location | 1 | 0.23257 | 0.23257 | 6.71 | 0.015 |
| Dog\*Operator | 2 | 0.26358 | 0.13179 | 3.80 | 0.034 |
| Dog\*Day | 2 | 0.51395 | 0.25698 | 7.42 | 0.002 |
| Location\*Operator | 2 | 0.50805 | 0.25403 | 7.33 | 0.003 |
| Location\*Day | 2 | 0.06244 | 0.03122 | 0.90 | 0.417 |
| Error | 30 | 1.03922 | 0.03464 |  |  |
| Total | 47 | 4.44896 |  |  |  |

* The factor Treat is most significant, with a p-value of 0.000, followed by the interaction between Dog and Treat.
* The factor Operator is marginally significant.
* Factor effects for Dog, Location, Day, and Location\*Day are not significant.

**5.5 ANOVA Assumption Validation and Residual Analysis**

ANOVA analysis requires validation of key assumptions. Those assumptions are:

1. Residuals are normally distributed
2. Run-order independence
3. “No outliers” (5% of normally distributed points lie beyond 2 standard deviations from the mean)
4. Homogeneity of Variance for residuals



The residuals are normally distributed, having a p-value of 0.785 for the Anderson-Darling normality test. Validating assumption (1).

There is no recognizable pattern in the “Versus order” plot in figure 5.5.1, which is sufficient justification to validate assumption (2).

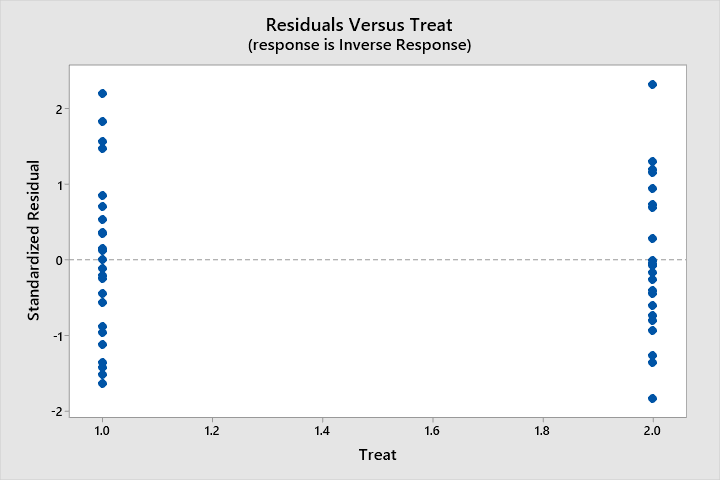
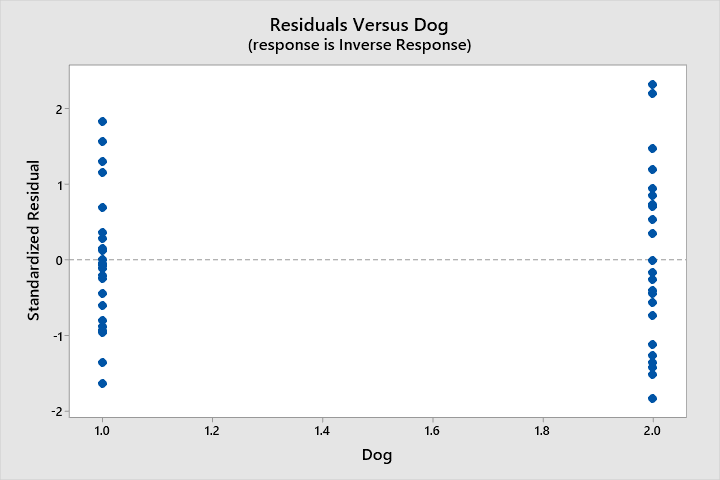
There are 2 “potential outliers” having standardized residuals of 2.20 and 2.32. It is expected to find 5% of normally distributed values outside 2 standard deviations from the mean. Given a sample size of 48, these 2 values are approximately 4% of the data. Assumption (3) is validated.

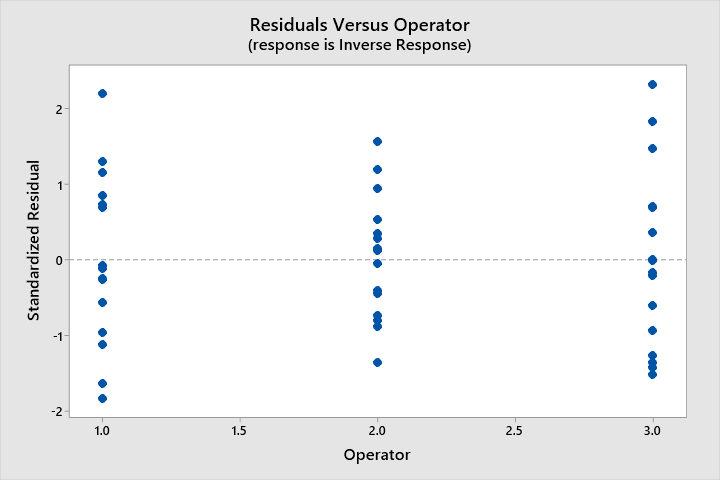
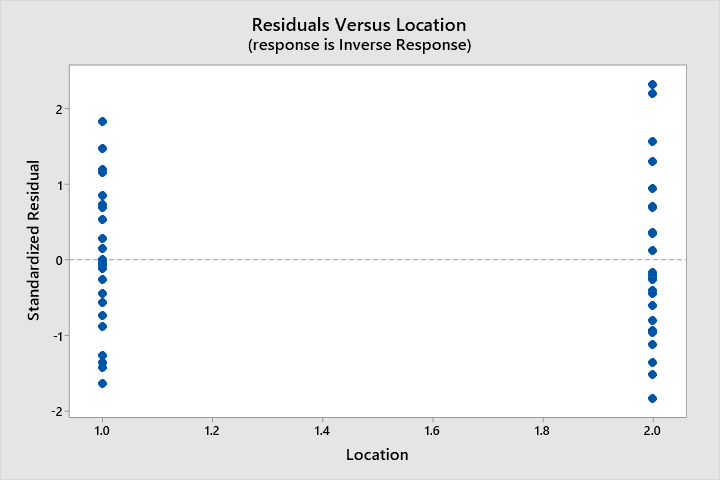
**Table 5.5.1: Fits and Diagnostics for Unusual Observations**

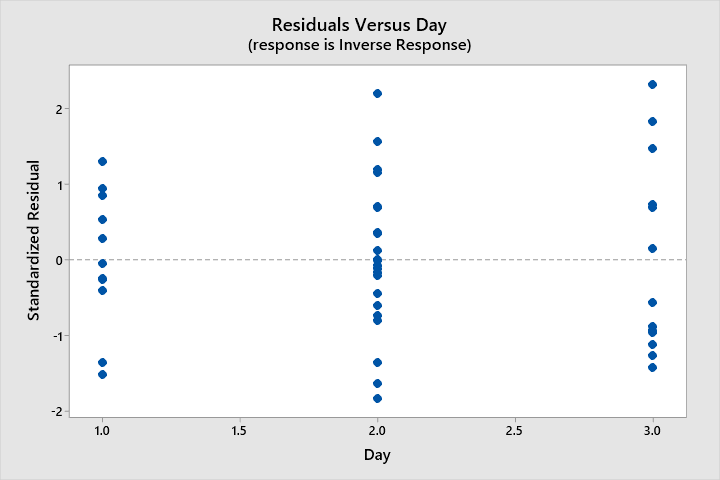
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Obs** | **Inverse Response** | **Fit** | **Resid** | **Std Resid** |  |
| 28 | 1.408 | 1.077 | 0.332 | 2.20 | R |
| 39 | 0.521 | 0.228 | 0.292 | 2.32 | R |

The residuals appear to have a near constant variance about the expected value in the versus fits plot in figure 5.5.1. The residual plots and Levene’s test of homogeneity of variance table below do not indicate any major homogeneity of variance concerns. This is sufficient justification to validate assumption (4).

**Figure 5.5.2: Residual Plots by Factor**







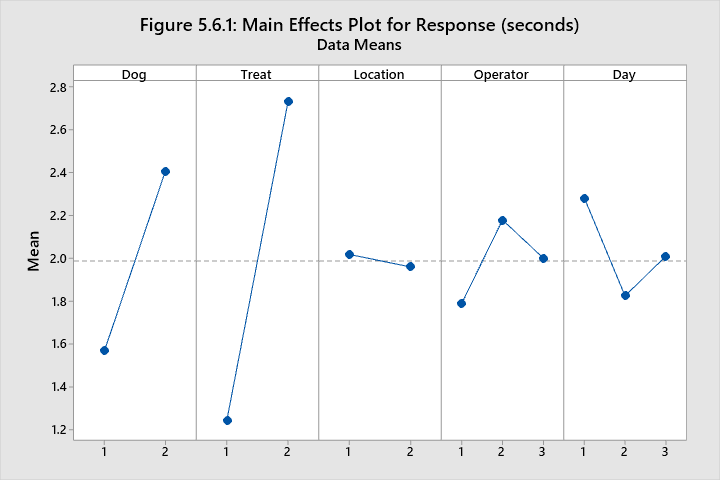
**Table 5.5.2: Levene’s Test of HOV Results by Factor**

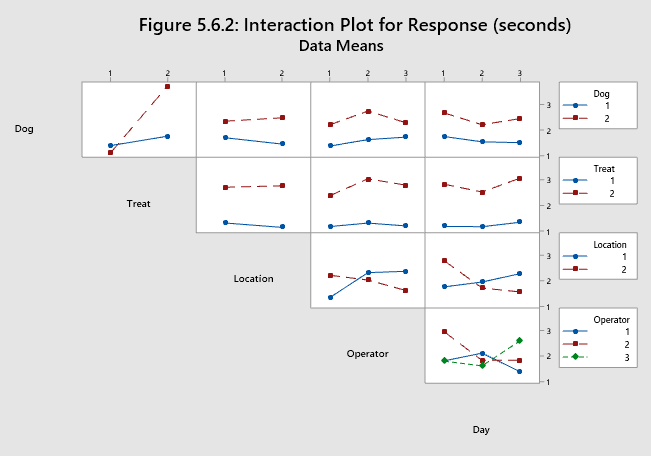
|  |  |  |
| --- | --- | --- |
| **Factor** | **Test Statistic** | **P-Value** |
| Dog | 1.82 | 0.184 |
| Treat | 0.30 | 0.588 |
| Location | 0.44 | 0.509 |
| Operator | 0.88 | 0.421 |
| Day | 1.29 | 0.285 |

**5.6 Main Factor Effects and Interactions**

The factor and factor interaction effects can be visualized in figures 5.6.1 and 5.6.2 below. The Treat effect appears to have the largest main effect, and the interaction between Dog and Treat is especially interesting. The most significant conclusions from these plots are:

* Overall, using a treat to assist with the command “sit” appeared to decrease the response time by 1.5 seconds on average.
* While both dogs responded faster on average to the command when a treat was involved, dog 2 was especially slower when there was no treat, while it made little difference for dog 1.
* Dog 1 responded a bit faster outside than inside while dog 2 responded a bit slower outside compared to inside.





**5.7 Tukey Pairwise Comparisons for Significant Factors and Interactions with α = 0.05**

*\*Means that do not share a letter are significantly different.*

**Table 5.7.1: Tukey Pairwise Comparisons: Treat**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Treat** | **N** | **Mean** | **Grouping** | |
| 1 | 24 | 0.917400 | A |  |
| 2 | 24 | 0.495485 |  | B |

**Table 5.7.2: Tukey Pairwise Comparisons: Dog\*Operator**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dog\*Operator** | **N** | **Mean** | **Grouping** | |
| 2 1 | 8 | 0.822997 | A |  |
| 1 2 | 8 | 0.772199 | A | B |
| 2 3 | 8 | 0.761079 | A |  |
| 1 1 | 8 | 0.732961 | A | B |
| 1 3 | 8 | 0.720530 | A | B |
| 2 2 | 8 | 0.428889 |  | B |

**Table 5.7.3: Tukey Pairwise Comparisons: Dog\*Treat**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dog\*Treat** | **N** | **Mean** | **Grouping** | | |
| 2 1 | 12 | 0.993232 | A |  |  |
| 1 1 | 12 | 0.841569 | A | B |  |
| 1 2 | 12 | 0.642224 |  | B |  |
| 2 2 | 12 | 0.348745 |  |  | C |

**Table 5.7.4: Tukey Pairwise Comparisons: Dog\*Day**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dog\*Day** | **N** | **Mean** | **Grouping** | |
| 2 2 | 10 | 0.900886 | A |  |
| 1 1 | 5 | 0.858364 | A | B |
| 1 3 | 6 | 0.782810 | A | B |
| 2 1 | 7 | 0.614970 | A | B |
| 1 2 | 13 | 0.584515 |  | B |
| 2 3 | 7 | 0.497109 |  | B |

**Table 5.7.5: Tukey Pairwise Comparisons: Location\*Operator**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Location\*Operator** | **N** | **Mean** | **Grouping** | |
| 1 1 | 8 | 0.928271 | A |  |
| 2 2 | 8 | 0.800370 | A |  |
| 2 3 | 8 | 0.749919 | A |  |
| 1 3 | 8 | 0.731690 | A |  |
| 2 1 | 8 | 0.627686 | A | B |
| 1 2 | 8 | 0.400718 |  | B |

Minimizing the response time is desirable, therefore the best factor selection will require maximizing the transformed response. This is because as response increases, the inverse of response decreases, and vise-versa. If one desires to have the fastest response time for a command, according to this analysis:

* A treat should be used (Table 5.7.1)
* Dog 2 should be paired with operator 1 (Table 5.7.2)
* Dog 2 should be encouraged with a treat (redundant, Table 5.7.3)
* The day interaction cannot be selected for (Table 5.7.4), and
* Operator 1 should perform the experiment inside (5.7.5)

*Therefore, the optimum factor combination is (Dog, Treat, Location, Operator) = (2, 1, 1, 1).*

**Section 6: Final Conclusions**

Referring back to section 2.4 (Hypothesis Testing), the hypotheses to be accepted/rejected are below.

1. Holding a treat affects the mean time it takes for both dogs to obey the command.

**Reject** H0: μtreat = μwithout (in almost every case, the treat reduces response time)

1. The location which the experiment is run influences the response.

**Accept** H0: μinside = μoutside (the mean response time is the same inside and outside on average)

1. The ‘operator’ of the experiment influences the response.

**Accept** H0: μoperator1=μoperator2=μoperator3 (the mean response time is the same for operators on average)

The conclusions of this analysis are valid insofar as the experimental conditions are consistent. More data and experimentation would be needed to draw conclusions about the variability of dogs’ response time in general given these factors.

**Section** **6: References**

* + - 1. Factors affecting response of dogs to obedience instruction a field and experimental study **https://core.ac.uk/download/pdf/54294.pdf**
      2. Characteristics of Excitable Dog Behavior Based on Owners’ Report from a Self-Selected Study **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4810050/**

**Section 7: Self-Assessment**

