

# **Webflix Report on Minimizing User Browsing Time**

## **Executive Summary:**

In a recent survey of Webflix users who did not renew their subscriptions, 65% reported that the cause was due to choice-overload. To retain users and improve their viewing experience, Webflix has run experiments to minimize the amount of time users spend browsing. Webflix's team has found that to do so, the key values to optimize are: Tile Size, Match Score, Preview Length and Preview Type. Following multiple rounds of pre-experimentation with approximately 17,000 users, 990 user's browsing times were randomly selected. Using multiple linear models it was found that for the best browsing time, a match score of 72.88 and a teaser trailer preview of 75.02 seconds provides the most optimal browsing times, while tile size had a negligible impact in determining the optimal configuration.

## **Introduction:**

With a rise in alternative streaming services, Webflix has had to find ways to stand out from their competition through exclusive content deals and innovative ways to improve their user's experience. From a poll of current users and an exit survey of users who chose not to renew their subscriptions it was shown that 35% of existing users and 65% of past users said that they spent too much time browsing for a show to watch. This is known as decision paralysis or choice-overload, a phenomenon where users are overwhelmed with all the different options and eventually lose interest or decide to no longer watch a show.

To improve user experience Webflix identified four variables of the homepage: Tile Size, Match Score, Preview Length, and Preview Type that can be manipulated to minimize user browsing times. To find the optimal values, Webflix ran multiple experiments on sample user data to find the best models and methods to find the optimal values, settling on using a combination of factorial experiments to find the variables that are significant and response surface methodology to find the optimal values of each significant factor.

This report will cover the three phases of Webflix's experimental process used to identify the most optimal values and variables that will minimize a user's browsing time. The first phase is the initial screening phase where Webflix used a large set of randomly sampled user data to narrow down the breadth of the values in the experiment. The second phase consists primarily of Webflix testing its methodology on a more focussed range of data that was found in phase one. The third and final phase is where Webflix took the methodology and best practices from phases 1 and 2 and applied it on 990 observations of Webflix user data to find the optimal values for each variable to minimize browsing times.

## **Experimentation:**

### **Phase 1: Initial Screening**

In phase 1, before any experimentation began, the Webfix team identified four key variables of the Webfix homepage that they could change to optimize user browsing times: Preview Length, Preview Type, Tile Size and Match Score. Within the first phase, Webfix used a factorial experiment to examine the main effects of the different variables on users browsing times, resulting with figures 1, 2 and 3 being the resulting main effect plots for Preview Length, Match Score and Preview Type.

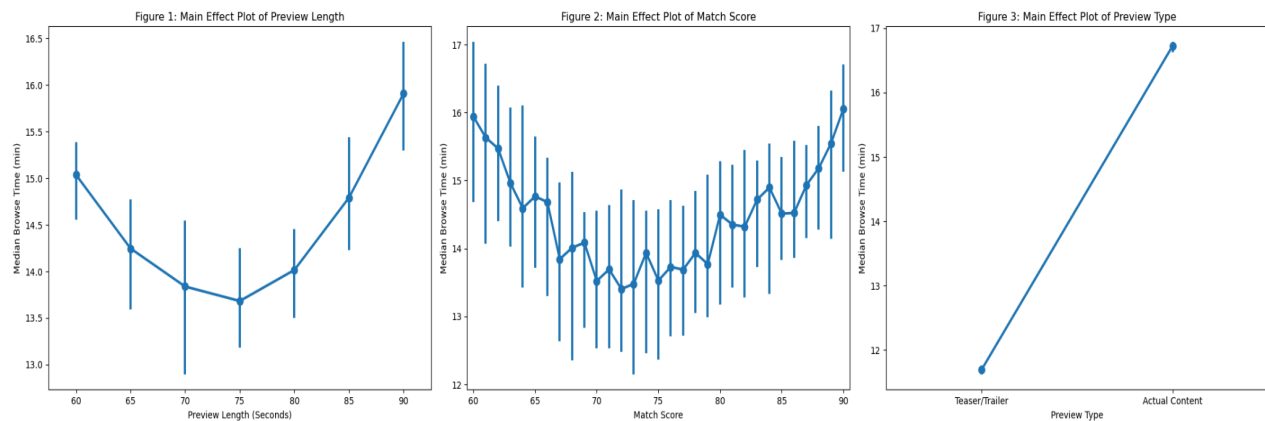


Figure 1 showing the main effect of preview length on user browse time, provided evidence to the Webfix team that the optimal preview length should be between 60 and 90. While figure 2 which examines the effect of match score, appears to show that the optimal value is between 60 and 90 as well. Figures 1 and 2 show that both match score and preview length appeared to have significant impacts on browsing time, with statistical testing in phase 2 used to prove this hypothesis. Figure 3, the preview type main effect plot, shows a significant impact based on the type of preview shown. When the Webfix team examined the effects of tile size on browsing time, it appeared to not be a significant predictor, appearing to have minimal effects. These preliminary findings are what allow Webfix to narrow the range of values and provide a base range for each variable in later phases, limiting the overall number of observations that are needed.

### **Phase 2: Factorial Design with Refined Ranges**

Next, the Webfix team refined the factor ranges to evaluate their impact on browsing time and to determine whether statistically significant results could be achieved. To ensure the experiment would have adequate statistical power, the team recalculated the required sample size, aiming for a statistical power of 95% and the 5% significance level. This was done using a root-finding approach based on Brent's optimization method, which determined the smallest sample size that would guarantee the desired power while controlling for Type I error. The calculation was based on the effect size and standard deviation, which were derived from the results of the initial phase. As a result, the required sample size per condition was found to be three. This sample size ensured sufficient power for the experiment, helping to guide the experimental design and the number of repetitions necessary for reliable conclusions.

The refined factor ranges included Preview Length and Match Score, both spanning from 60 to 90, and Tile Size ranging from 0.1 to 0.5. The team employed a factorial design to construct a series of hypothesis tests to assess the main effects and interactions of the factors. The hypothesis tests were structured as follows:

For main effect tests:

$H_0$ : The factor has a significant effect on the browsing time

$H_A$ : The factor has no significant effect on the browsing time

For interaction effect tests, including two-way, three-way interaction:

$H_0$ : There is no interaction effect between the factors on the browsing time

$H_A$ : There is a significant interaction effect between the factors on the browsing time

The testing was constructed using ANOVA analysis which allowed the team to identify and quantify the significance of factor contributions, both individually and in combination.

Model Name	F-statistic	p-value
Four-Way Interaction	0.458	0.499
Three-Way Interaction	1.017	0.398
Two-Way Interaction	160.2	0
Tile Size : Match Score	1.306	0.253
Tile Size : Preview Length	0.329	0.566
Tile Size : Preview Type	0.236	0.627
Match Score : Preview Length	958.7	0
Match Score : Preview Type	0.422	0.516
Preview Length : Preview Type	0.221	0.64
Tile Size	0.055	0.814
Match Score	0.009	0.92
Review Length	349.2	0
Review Type	18892.1	0

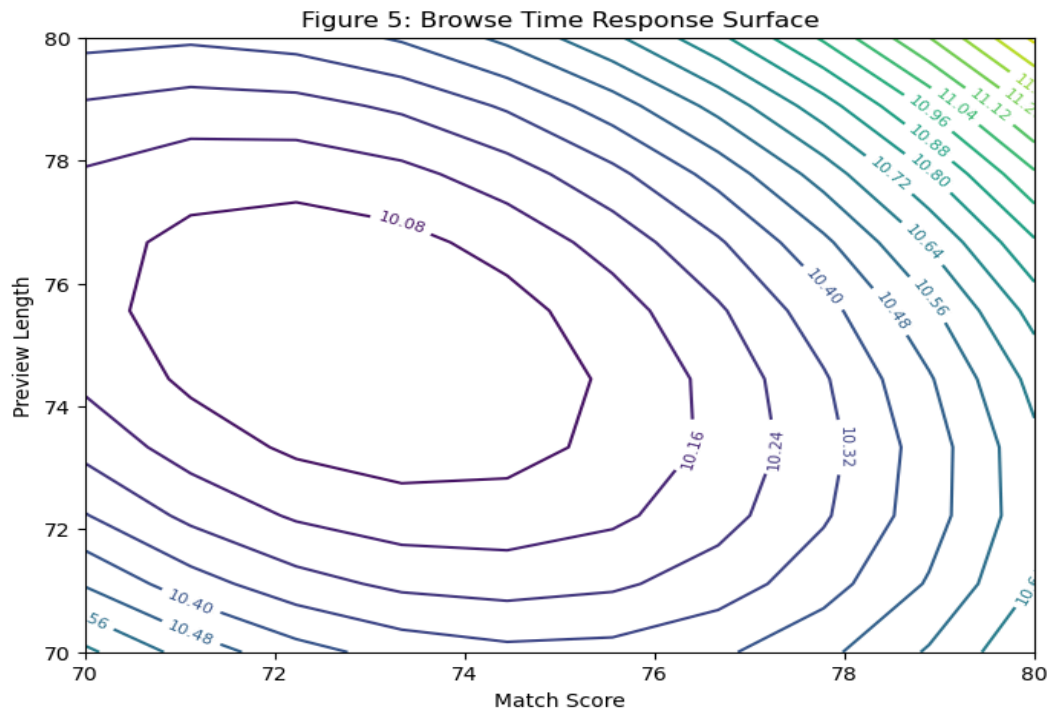
**Summary ANOVA Tests Table:** Each row in the table represents the results of an ANOVA (Analysis of Variance) test comparing a reduced model with a larger model to evaluate whether including additional terms significantly improves the model.

The summary table indicates that Preview Length, Preview Type were statistically significant main effects, while the interaction between Match Score and Preview Length were also statistically significant at the 5% significance level. These findings highlight the critical importance of these factors and their interactions, confirming the necessity of further narrowing the factor range. By leveraging the insights gained during this phase, the team was able to design more precise experiments such as refining the ranges for Preview Length, Match Score, and Preview Type to enable more targeted exploration in the next phase.

### **Phase 3: Further Refinement and RSM**

In the final phase, Response Surface Methodology (RSM) was implemented to fine-tune the optimal setting for the significant factors identified in Phase 2. Building upon the insights gained from previous phases, the team reconstructed a comprehensive experiment design matrix that focused on the most promising parameter ranges. For Match Score, values ranging from 70 to 80 were examined, while Preview Length was tested at 5-second increments between 70 and 80 seconds. Since Tile Size has minimal impact, the team decided to keep it at the default value of 0.2. Regarding Preview Type, Teaser Trailer had a much stronger effect compared to Actual Content. With that being said, 33 unique experimental conditions were created for testing.

A total of 990 user samples, with 30 units per condition, were randomly assigned across these conditions to ensure robust statistical validity. This larger sample size represented a significant rise from Phase 2 thereby providing greater statistical power and more reliable results. To find the optimal values of Preview Length and Match Score a two-factor central composite design was run in order to fit the second-order response surface model. The team believed that implementing this methodology would underscore the significant interaction between Match Score and Preview Length in driving user engagement.



As a result, Figure 5 illustrates that the stationary point for this second order model is located at a Preview Length of 75.02 seconds with the Match Score of 72.88. The estimated expected browsing time at this point is 10.00 minutes, with a 95% confidence interval of (9.88, 10.13).

## **Conclusion:**

Throughout the experimental process the main limitation that the Webflix team faced was a limit on the number of observations that could be gathered. In Webflix's attempts to mitigate this limitation, Webflix decided to capture as many possible conditions using an initial sample size of 3; before narrowing down the number of conditions and increasing the sample size to 30. Webflix believes that by going for breadth over depth in the experimental process allowed for Webflix to narrow down the ranges for possible optimal values for each variable in one process. Though Webflix believes this approach was optimal, the use of small sample sizes could result in the effects of some variables or interactions being exaggerated and showing larger or smaller effects.

Based on phases 1-3 the Webflix team was able to determine that the best way to improve user browsing times is to: move the algorithm's Match Score value rounded to 73, Preview Length to 75 seconds and change the Preview Types to showing Teaser Trailers. By doing so Webflix believes that they will be able to reduce user's browsing time by around 10 seconds, lowering the chances of decision paralysis and increasing Webflix's competitiveness. Moving forward Webflix will continue to work to improve the user experience by tracking the effects of the current proposed changes as well as looking for other aspects of the user experience that could be optimized.