

# Redesigning Skill-Based Metrics in Flash Flash Revolution

Wilson Cheung (WirryWoo)  
*Georgia Institute of Technology*

**Abstract:** The objective of this study is to explore methods to improve the scalability and standardization of skill-based metrics in Flash Flash Revolution. For stepfile difficulty, we propose a new method that combines ensemble tree-based techniques with a time series extrinsic regression model to capture local and global patterns in stepfiles using objectively defined time-sensitive features. Additionally, we introduce a framework that utilizes the expectation-maximization algorithm to determine a player's AAA equivalency rating and overall skill rating, providing a quantitative measure of their skill. These machine learning models have been implemented in ACubed, a Python module designed to continuously update the underlying formula for skill-based metrics, in order to enhance the competitive experience for all players on FFR.

## Contents

<b>1. Introduction and Motivation</b>	<b>2</b>
1.1. Overview of FFR's Skill-Based Metrics and Their Shortcomings . . . . .	2
1.2. Modeling Disadvantages of Using Meta-Established Terminology . . . . .	5
1.3. ACubed: Revamped Framework for FFR's Ranking System . . . . .	9
<b>2. Stepfile Difficulty Model</b>	<b>10</b>
2.1. Contested Difficulty Spreadsheet Analysis . . . . .	11
<b>3. AAA Equivalency Model</b>	<b>12</b>
<b>4. Player Skill Rating Model</b>	<b>12</b>
<b>5. Future Work</b>	<b>13</b>
<b>6. Technical Implementation</b>	<b>15</b>
<b>7. Conclusion and Acknowledgements</b>	<b>16</b>
<b>A. Appendix</b>	<b>18</b>

# 1. Introduction and Motivation

Flash Flash Revolution (FFR) is a renowned Flash-based vertical scrolling rhythm game that was originally developed by **Synthlight** in 2002. Similar to its predecessor Dance Dance Revolution (DDR), players are presented with a sequence of arrows, referred to as *stepfiles*, which are meticulously calibrated and synchronized to the background music. The main objective is to maximize the raw score by pressing the corresponding keys within a precise timing window of less than one-quarter of a second, ensuring a perfect alignment between the arrows in motion and the static receptors. Following the deprecation of Adobe Flash Player in 2020, FFR is now accessible through RCubed, an open-source engine developed and actively maintained by **Velocity** [Tea].

The player base of FFR maintains a consistently robust presence, with several thousand users actively engaging in gameplay every month. Among this devoted community are competitive players hailing from online rhythm games such as osu!mania and Etterna, with a notable number participating in official tournaments each year. To ensure a positive experience for these players, many systems have been created to objectively measure stepfile difficulty, fairly reward scores based on playing performance, and determine ratings that accurately reflect the player’s skill. Although these systems are widely adopted in the community, their design has sparked considerable concerns regarding the reliability and scalability of their skill measurement.

Before delving into these areas of discussion, it is important to acknowledge that there will be many assumptions present throughout this analysis. We will provide sufficient justification for any assumptions made, and it is important to note that the proposed solutions in this paper are not absolute and definitive. If there is any disagreement with any identified assumptions, it is encouraged for individuals to develop an alternative modeling approach based on their preferred assumptions.

## 1.1. Overview of FFR’s Skill-Based Metrics and Their Shortcomings

### 1.1.1. Stepfile Difficulty

Every stepfile possesses intrinsic levels of challenge concerning the player’s execution due to their distinct structural characteristics. However, quantifying this measure continues to be a challenging question. In the initial stages of its development, the measurement of stepfile difficulty in FFR was based on a scale ranging from 1 to 12, where 1 represented the easiest and 12 represented the most difficult level. However, in 2012, **nois-or-e**, **One Winged Angel**, and **stavie33** improved upon this method by introducing a more granular scale ranging from 1 to 120. This update was announced on the forums [nAS] and has since been adopted as the current difficulty scale in use.

To assess the difficulty of new stepfiles in the engine, a team of difficulty consultants within the community assigns a numerical rating on an integer scale of 1 to 120, taking into consideration factors such as feedback from other players and the level of challenge posed by existing stepfiles of similar difficulty ratings. However, this process raises the following concerns:

- By solely relying on pairwise stepfile comparisons to determine the difficulty rating of an unrated stepfile, other assigned ratings are not considered during comparison, resulting in an inconsistent set of criteria established across stepfiles to define difficulty.
- Proposed difficulty ratings crowdsourced from the vocal minority are inherently subject to self-selection bias by the current process’s mechanism design. Moreover, manual stepfile difficulty assignment for easier files is generally less accurate due to higher skill bias in the vocal minority.
- There is an inability to predict difficulty at any rate not equal to 1.0x, which has recently gained popularity for competitive play within the community.

Additionally, this current method of determining the difficulty of stepfiles is a manually daunting task, prompting many developers to seek algorithmic methods to measure stepfile difficulty automatically. With the growing advancements in machine learning and artificial intelligence, these methods have become increasingly popular for creating stepfile difficulty models. Recently in 2023, **Trumpet63** successfully developed an API wrapper incorporating a trained GRU model with a mean-squared error of 11 [Trua], to help estimate the difficulty of new stepfiles in the FFR Batch Submission page.

Although **Trumpet63**'s model has primarily received positive sentiment from the FFR community, these endeavors continue to face criticism, as there is a general lack of confidence in the model's ability to accurately capture the multifaceted nature of this issue. This skepticism is further compounded by the lack of transparency in the currently used model, which gives rise to the following concerns:

- There are numerous interpretations of stepfile difficulty, making it challenging to establish a standardized definition that is universally accepted by the community.
- There is a subjective component inherent in the concept of difficulty that needs to be decoupled from the overall dataset to appropriately perform model training and inference.
- There is uncertainty on how stepfile patterns (e.g. jacks, jumpstream, etc.) can be appropriately defined and integrated as reliable features in model training.

Given the progress made in measuring stepfile difficulty and its widespread use as a means of facilitating discussions on player's proficiency, it is evident that this measure holds great significance in the FFR community. Therefore, it is imperative to establish a definitive and thorough understanding of stepfile difficulty, in order to mitigate any further challenges in defining other skill-based metrics.

### 1.1.2. AAA Equivalency

The *AAA equivalency* was developed to determine the level of difficulty at which a player's performance on a specific stepfile can result in the highest attainable score (AAA). Particularly, the equivalence is determined by comparing the number of "raw goods" obtained on a given difficulty level to scoring a AAA on another difficulty level. The initial formula defining this metric was proposed by **Trumpet63** after conducting a graphical analysis of player performance during the 10<sup>th</sup> Official FFR Tournament [Trub]. However, as official tournaments have continued, the formula has been continuously updated and refined based on new data. These adjustments have resulted in the current version of the formula [Pra], which is outlined below.

$$AAA_{eq} = (D + \alpha) \left( \frac{\delta - NGC \cdot \lambda}{\delta} \right)^{1/\beta} - \alpha$$

$$NGC = \text{Goods} + 1.8 \cdot \text{Averages} + 2.4 \cdot \text{Misses} + 0.2 \cdot \text{Boos}$$

$$D = \text{Difficulty}$$

$$\delta = a_0 + a_1 \cdot D + a_2 \cdot D^2 + a_3 \cdot D^3 + a_4 \cdot D^4$$

where

$$a_0 = 17678803623.9633$$

$$a_1 = 733763392.922176$$

$$a_2 = 28163834.4879901$$

$$a_3 = -434698.513947563$$

$$a_4 = 3060.24243867853$$

and

$$\begin{aligned}\lambda &= 18206628.7286425 \\ \alpha &= 9.9750396740034 \\ \beta &= 0.0193296437339205\end{aligned}$$

Although the notion of AAA equivalency is widely accepted and acknowledged in the community, there are several concerns with its implementation. These concerns can be partly attributed to the lack of a consistent and accurate definition of stepfile difficulty.

- Certain stepfiles have been identified to be more conducive to maximizing AAA equivalency, leading to unwarranted inflation when measuring the skill rating of a player.
- Players who attain impressive results with a high number of raw goods may receive inaccurate and undeserved ratings, particularly for many of the most difficult stepfiles in FFR.
- The current AAA equivalency formula poses challenges in its interpretation, causing individuals to question the rating they have received.

Several solutions have been proposed on Discord to address the issues identified with the AAA equivalency metric. Proposed by **Lights**, a potential solution that shows promise suggests incorporating additional weighting in the formula to prevent players from exploiting the metric and inflate their skill ratings, while also accurately recognizing AAA equivalency for players who score above the current threshold of 28 raw goods on some of the most difficult stepfiles in FFR. However, these solutions are still in the process of being investigated and no significant progress has been made thus far.

### 1.1.3. Skill Rating

The *Skill Rating* metric, which is determined by analyzing players' top-performing scores across their entire play record, serves as the primary basis for ranking players in FFR. Introduced by **Trumpet63**, **PrawnSkunk**, and **Silvuh**, this ranking metric is initially defined by computing an exponentially weighted average of a player's top  $N$  scores [Pra]. Particularly, the weighting mechanism that was rolled out in its initial release defines the weights  $w := w(k, r)$  using the following formula for all ranks  $r$  satisfying  $r \in \{1, 2, \dots, N\}$ :

$$w = \frac{(e^k - 1)(e^{k(r-1)} - e^{kN})}{(e^{kN} - 1) - Ne^{kN}(e^k - 1)}$$

where  $k = -\frac{1}{4}$  and  $N = 15$ .

Under this configuration, note that over 75% of a player's skill rating is determined by the Top 5 scores. With the growing amount of stepfiles in FFR, there is an increase in the likelihood of a player obtaining a high score through chance, ultimately resulting in an incorrect representation of their true skill under this particular weighting mechanism. Moreover, relying on performance data indicative of a player's skill ceiling has been known to result in incorrect official tournament placements each year, thus creating an unfair disadvantage for certain players and diminishing their chances of winning.

To address some of these challenges, a new approach called Seasonal Skill Rating was implemented in conjunction with the established Skill Rating model in 2021 [PXW]. This system was specifically designed to monitor player performance in three-month intervals each year, providing more current data for use in determining official tournament placements. Developed by **Prawkskunk**, **XJ-9**, and **WorryWoo**, the weighting mechanism used in Seasonal Skill Rating calculates a linearly weighted average of a player's top  $N$  scores using the formula provided below. For all ranks  $r$  satisfying  $r \in \{1, 2, \dots, N\}$ , the weight  $w := w(r)$  is computed as follows:

$$w = \frac{N - r + 1}{\frac{1}{2}N(N + 1)}$$

For the initial implementation of the Seasonal Skill Rating,  $N = 50$  was deliberately selected to evaluate its performance in comparison to the established Skill Rating metric. Based on favorable reviews following the first season’s conclusion, the weighting mechanism of the Skill Rating model was revised to incorporate the new formula, which has since remained the primary approach for determining skill rating.

Despite the improvements, there are still several limitations in the Skill Rating model that necessitate further investigation:

- With the continuous addition of new stepfiles in FFR, players may have more opportunities to exploit  $N$  particular stepfiles to achieve fluke scores. Conversely, for players with incomplete or unrepresentative score data, this may not accurately reflect their skill level. This raises the question of whether  $N$  should be adjusted by the number of stepfiles played.
- The Seasonal Skill Rating only considers a player’s top scores on stepfiles, but there is a need for a more comprehensive measure of all scores within a reasonable range from the player’s Skill Rating, to improve tournament placements.

Regrettably, due to insufficient availability of data, it is improbable that we will uncover resolutions for a number of the aforementioned issues. Nonetheless, these potential solutions will be outlined in the report to offer additional understanding on the necessary steps to tackle these problems.

## 1.2. Modeling Disadvantages of Using Meta-Established Terminology

In this section, we shift our focus to a contentious topic and explore the drawbacks of using meta stepfile features and raw goods in our machine learning models designed to improve the overall ranking system.

### 1.2.1. Stepfile Patterns

During gameplay, players observe similar patterns within numerous stepfiles and use these patterns to centralize discussions about strategies to optimally hit them. These patterns, referred to as *meta stepfile features* in this report, are also currently used to draw comparisons between two different stepfiles to determine their level of difficulty. These include but are not limited to the following:

- A *jump* is a common pattern in which players must simultaneously hit two notes. The term originated in DDR, as executing a jump in the game requires players to physically jump to successfully hit the notes.
- *Jumpgluts* comprises a rapid succession of jumps occurring within a short timeframe.
- *Jumpchains* are sequences that alternate between jumps and a single connecting arrow. The connecting arrow is required to be the same between the two neighboring jumps.
- A *hand* is a common pattern in which players must hit three notes simultaneously. The term “hands” denotes a type of jump on a dance pad involving three notes, which necessitates using both feet and a hand for execution.
- A *quad* is another common pattern in which players must simultaneously hit four notes.

- *Grace notes* are defined as concise, rapidly executed notes that lead into a subsequent note, representing sustained noise. During gameplay, these can often be hit as a jump due to their close timing proximity.
- *Trills* are alternating patterns of two arrows, with *one-handed trills* requiring players to continuously alternate between the two arrows with only one hand.
- *Jumptrills* refer to trills characterized by two alternating jumps that occur consecutively without any note overlap.
- *Jacks* are arrow sequences that involve hitting the same arrow three or more times in a row. When it is only required to execute two consecutive hits, it is classified as a *minijack*.
- *Jumpjacks* are a type of jacks in which the same jump is executed three or more times, whereas *minijumpjacks* involve executing the same jump twice.
- *Chordjacks* refers to a collection of patterns that incorporate different combinations of jacks and chords. This term has gained prominence within the meta and has become increasingly prevalent in recent times.
- *Streams* are uninterrupted successions of single notes, typically performed at quick tempos and may vary in note quantization. A *roll* is a particular type of stream that has a defined sequence of patterns.
- *Jumpstreams* and *handstreams* are types of streams characterized by an increased level of complexity due to the inclusion of jumps and hands, respectively.
- *Triplets* refer to a broad category of isolated patterns that consist of three individual notes played in succession, while *gallops* are characterized by a similar structure consisting of two successive notes.
- *Running men* features a recurrent anchor arrow within a stream, with the remaining notes alternating between two or three other arrows. Note that in the case of a single alternate note, it is classified as a trill.
- *Staircases* refer to patterns that evoke a sense of ascent or descent, achieved by traversing the arrow keys on the keyboard in a left-to-right or right-to-left sequence.
- *Bursts* refer to intricate arrangements of notes that occur in quick succession, often without adhering to a specific structure in their sequencing.
- *Polyrhythms* refer to the occurrence of two distinct rhythms played simultaneously, resulting in a complex and disorienting pattern.

However, there are numerous obstacles that arise when attempting to apply definitions of meta patterns from a stepfile to different rates. For instance, at very high rates, rolls compressed in a condensed time period (as demonstrated in Figure 2, assuming that the note spacing in Figure 1 is significantly larger than the allocated perfect window) are often treated as jumptrills. This technique, known as "jumptrilling," is commonly used in competitive play to maximize scores in FFR. As this skill is easily accessible for many players, it is natural for them to equate it with the overall difficulty of the stepfile, potentially leading to a perception that the file is easier than initially thought.

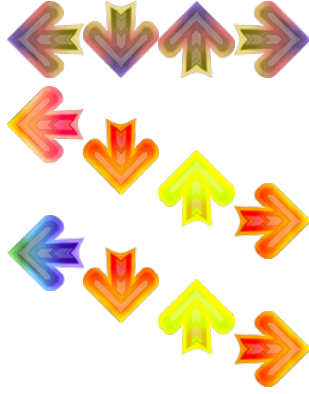


Figure 1: Non-“jumptrillable” rolls under sufficiently slow scroll speed in upscroll direction.

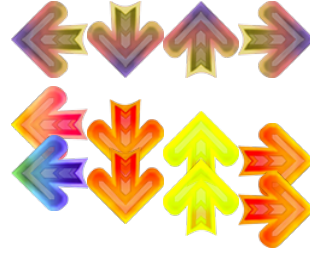


Figure 2: “Jumptrillable” rolls under sufficiently fast scroll speed in upscroll direction.

On the contrary, the execution of jack patterns is commonly recognized as a highly personalized skill, with lower-ranked players often demonstrating greater proficiency in these patterns compared to their higher-ranked counterparts. As a result, certain players may perceive hitting a single note at a consistent tempo as a sequence of jacks (Figure 4), while others may interpret it as individual notes (Figure 3), particularly at different song speeds. This raises the question of what factors must be present in order to distinguish single notes that are the same in orientation from jack patterns in stepfiles. Conflicting opinions on this matter, coupled with the inherent difficulty of executing jack patterns for most players, have resulted in varying perspectives among community members regarding the overall challenge posed by these patterns.

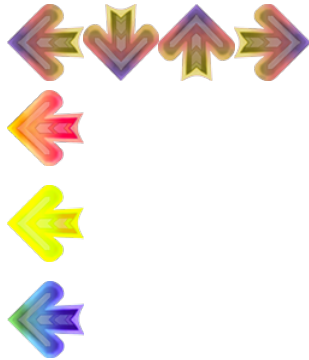


Figure 3: Singles under sufficiently slow scroll speed in upscroll direction.

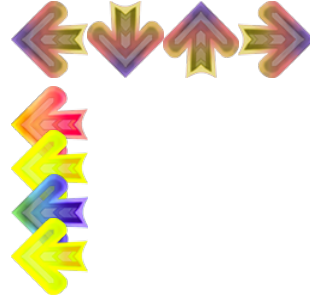


Figure 4: Jacks under sufficiently fast scroll speed in upscroll direction.

These examples illustrate the lack of a widespread consensus on how to differentiate between meta patterns across rates, creating additional challenges when attempting to create a machine learning model for measuring stepfile difficulty. In order for such a model to be successful, these criteria must be accurately defined and widely accepted within the community. Unfortunately, the lack of agreement

on this issue has impeded progress in this area and makes it unlikely that a resolution will be reached in the near future.

To address this issue, we suggest defining *objective stepfile features* that conform to a widely accepted standard of difficulty, regardless of the player’s level of skill and prior experience with the game. These features will be identified and further discussed in Section [TODO].

### 1.2.2. Raw Goods

*Raw Goods* is designed to translate all imperfect judgments into a standardized score to better assess the player’s performance on a given stepfile. This approach is based on FFR’s implementation of raw scoring, which was adopted as an alternative to combo scoring to discourage players from being rewarded for mindlessly mashing keys while playing. The point system for raw scoring is summarized below:

Judgement $j$	Score $s(j)$
Perfect	50
Good	25
Average	5
Miss	−10
Boo	−5

and raw goods  $r(j)$  is computed as follows:

$$r(j) = \begin{cases} \frac{s(j) - s(\text{Perfect})}{s(\text{Good}) - s(\text{Perfect})} & \text{if } j \neq \text{Boo}; \\ \frac{s(\text{Boo})}{s(\text{Good}) - s(\text{Perfect})} & \text{otherwise.} \end{cases}$$

This yields the following weights per note:

Judgement $j$	Raw Goods $r(j)$
Perfect	0
Good	1
Average	1.8
Miss	2.4
Boo	0.2

The aforementioned raw good values are subsequently aggregated according to the player’s inputs against the notes present in the stepfile. These values, in conjunction with the stepfile difficulty, are currently used to quantify a player’s performance using the AAA equivalency formula. Particularly, there exists an implicit, and potentially inaccurate, assumption that all notes possess equal significance in gauging a player’s performance. Further research is necessary to determine the potential to improve player skill measurement by implementing a weighted representation system for each note, where weights are defined concerning its level of difficulty within a stepfile.

Furthermore, the ranking system does not account for different note counts when evaluating the performance of players scoring identical raw good counts in two different stepfiles. This may result in inconsistent performance measurements across stepfiles, which could lead to unreliable assessments of an individual’s skill level. Consequentially, some stepfiles are also perceived to be more difficult to single-digit good (SDG) and AAA than others, causing the definition of difficulty to become unnecessarily conflated.

Based on the identified limitations of using raw goods, we will identify *score features* to improve accuracy in assessing the player’s performance on the stepfile. This includes defining an adjusted



variant of raw goods independent from note count and as previously mentioned, utilizing standard classification metrics from machine learning to measure performance based on the quantity of Misses and Boos. More details are to be covered in Section [TODO].

### 1.3. ACubed: Revamped Framework for FFR's Ranking System

Following a discussion on the limitations of utilizing commonly accepted terminology within the FFR community and an overview of the deficiencies in FFR's current rating system, this paper presents ACubed, a revamped ranking system designed to increase scalability, maintain accurate skill and difficulty measurement, and minimize required manual effort currently instilled in FFR's defined processes. The name ACubed has been chosen for the following reasons:

- The competitive culture within the FFR community focuses heavily on AAA ratings, making the name of this new system particularly relevant.
- As introduced in Chapter 2, the system being introduced will potentially be the third major update to FFR's difficulty system.
- ACubed is proposed as an enhancement to FFR's ranking system in spirit with RCubed being launched as a significant improvement to the game engine.

The summary of proposed changes is visually represented in the diagram shown in Figure 5.

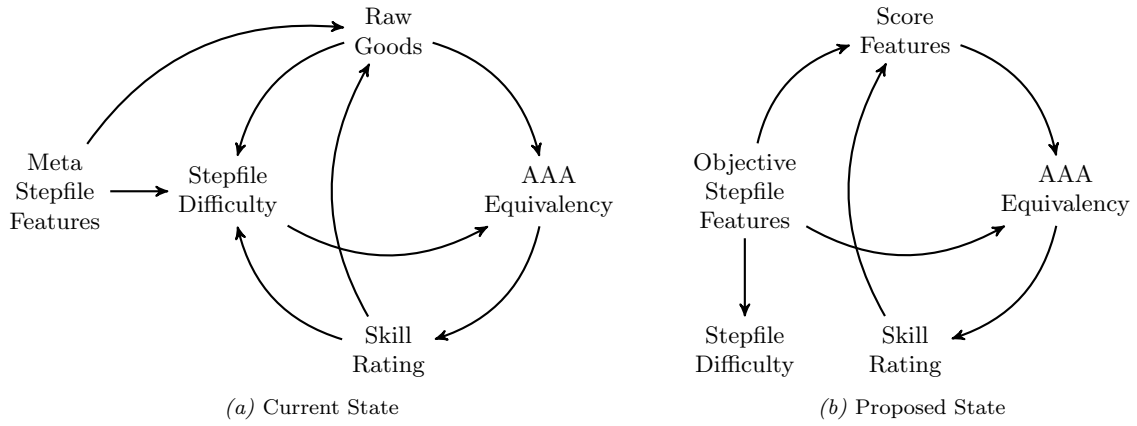


Figure 5: Dependency graphs differentiating the current and proposed states of skill-based FFR metrics.

From the diagram above, the following changes are recommended:

- The stepfile difficulty model will be separated from player performance metrics such as skill rating and raw goods to minimize player subjectivity.
- Objective stepfile features will be utilized instead of relying on patterns established by the community to promote consistent terminology.
- Raw goods will be replaced with score features to more accurately track player performance.
- The AAA equivalency formula will no longer involve stepfile difficulty, but instead will directly incorporate objective stepfile features.

Additionally, it is important to note that the Proposed State, as illustrated in Figure 5b, exhibits a circular interdependence between Score Features, AAA Equivalency, and Skill Rating. A well-designed system should be able to:

- Accurately assess a player’s performance on a stepfile by utilizing score features derived from a predetermined skill rating metric.
- Calculate AAA Equivalency without making the assumption of uniform distribution of notes within the stepfile and also take into account performance metrics that are more focused on the player.
- Continuously adjust a player’s skill rating without distorting their true skill level, which may be influenced by factors such as the number of stepfiles played.

The implementation of this study necessitates the development of a framework for training machine learning models using the expectation-maximization algorithm. Within this framework, the models are optimized concerning their corresponding dependent skill-based metrics based on a predetermined convergence criterion. Further details on this approach will be discussed in Section [TODO], with detailed steps for the technical implementation outlined in Section [TODO].

In Section 2, we aim to establish a clear definition of stepfile difficulty by addressing the inherent subjectivity of this concept. We will examine various existing interpretations and propose a methodology for estimating difficulty based on these definitions. This process includes creating a data stratification strategy to incorporate player feedback from `#difficulty-discussion` in FFR’s Official Discord channel within our train-test split, identifying a set of objectively defined features that align with specific gameplay properties, and devising a framework for training and evaluating a time series extrinsic regression model. This model will be used to replicate experiments under varying hyperparameter configurations and considered definitions.

Our definition of AAA equivalency is consistent with the interpretation of the community. However, in Section [TODO], we propose improvements to the current formula by defining an adjusted raw good count concerning features identified for our stepfile difficulty model. Additionally, we introduce precision and recall as performance-indicative metrics within the realm of competitive gameplay by defining the “Boo” judge accuracy label as false positives and the “Miss” judge accuracy label as false negatives.

In Section [TODO], we propose a study advocating for the incorporation of outlier detection techniques to eliminate unrepresentative AAA equivalency scores. We also present a framework for adjusting the value of  $N$  depending on the number of stepfiles played. Furthermore, we suggest incorporating Seasonal Skill Ratings to contextualize players’ abilities beyond their top-performing scores and evaluate their performance in comparison to Skill Ratings.

## 2. Stepfile Difficulty Model

**Plan:**

- Untangle objective definitions of difficulty from its inherent subjectivity. Identify that the goal of creating this machine learning model is to maximize the objective representation of difficulty in the overall model.
- Clarify the definition of difficulty by introducing judge window sizes. Also dissociate processing difficulty from execution difficulty and focus paper on execution. Add visuals:
  - Difficulty to FC: 235 ms

- Difficulty to Clean FC: 201 ms
- Difficulty to Maximize Score: 153.9 ms
- Difficulty to Maximize Score under MS based scoring: 117.5 ms
- Difficulty to AAA: 100 ms
- Difficulty to AAAA: 34 ms
- Address concerns about inaccurate difficulty measurements current in game and the non-existing idea of "true" difficulty using crowdsourcing papers [here](#) and [here](#).
- Introduce Contested Difficulty Sheet and define train-test split using KDE estimations of proposed difficulties
- Introduce objective features VerticalDensity and HorizontalDensity using pen-tapping as an example. Also mention chart augmentation considerations:
  - Mirror: Indirectly accounted in VerticalDensity's definition
  - Offset: Indirectly accounted in preprocessing of FFR's API response.
  - Colors: Categorized under "processing difficulty"; users can alter this.
  - Scroll Speed Mod: Categorized under "processing difficulty"; users can alter this.
- Dissociate local features from global features. Song length is a global feature, subsequence extraction is a local feature. Mention how ensembling works with the consideration of "difficulty spikes" in many charts.
- Introduce model architecture and implementation. Consider regressing over the target variable transformations of difficulty due to domain knowledge of difficulty.
- Run experiments with proposed model, share results, and identify method to measure the quality of the predictions.
- Extend proposed model to account for uprated and downrated stepfiles, run experiments and compare model performance with already tracked uprated scores.

## 2.1. Contested Difficulty Spreadsheet Analysis

The Contested Difficulty Spreadsheet, managed by Zlyice, is a collaborative Google Sheets document utilized for tracking proposed changes to stepfile difficulty measurements based on input from users. It is presumed that the individual suggesting a difficulty measurement possesses a degree of knowledge regarding the factors that warrant such a rating for the stepfile. While it must be acknowledged that all proposed measurements are inherently vulnerable to volunteer bias, the process of crowd-sourcing these partial inputs serves to reduce the impact of individual bias on any given stepfile. This approach aims to align the definition of difficulty with the perspectives of the vocal minority within the wider community.

In order to align our model scoring with the prevailing consensus on the quantification of stepfile difficulty within the community, we conduct a thorough analysis of the proposed difficulty levels over the period spanning from 2022 to the present date. This analysis serves as the basis for our development of a train-test split strategy, in which the distribution of proposed difficulties is utilized to allocate our data into appropriately representative training, validation, and test sets.

### 3. AAA Equivalency Model

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam fermentum sapien a felis mattis fermentum. Nunc laoreet eros in malesuada ultricies. Sed purus libero, blandit vitae convallis eu, lacinia ut justo. Curabitur est mauris, tincidunt in lacinia et, varius nec tortor. Nunc eu turpis feugiat, porta diam non, tincidunt lorem. Sed pulvinar, metus a volutpat malesuada, urna quam tempor nisi, nec cursus urna dolor vitae felis. Pellentesque posuere neque sapien, quis sagittis neque luctus quis. Quisque ullamcorper dignissim suscipit. Maecenas pharetra, nisi a pretium placerat, ante est posuere mauris, nec auctor mi diam in nisl. Pellentesque ligula magna, volutpat sed nunc vitae, auctor gravida arcu. Maecenas sit amet lacus justo.

In eu finibus purus. Etiam accumsan erat justo, eu efficitur elit faucibus non. In faucibus vitae nisl dapibus ullamcorper. In fermentum leo eu purus rutrum suscipit. Pellentesque dictum efficitur purus, eu consequat metus pretium a. Vivamus feugiat urna ac accumsan suscipit. Praesent pretium sem risus, a congue felis placerat sed. Maecenas in nisi auctor, aliquet tortor ut, posuere justo. Cras lacinia bibendum elit, ut cursus turpis eleifend vitae. Nunc tempus libero quis suscipit scelerisque. Curabitur aliquet tempor facilisis. Suspendisse auctor, arcu non dictum luctus, sem tortor varius lorem, non commodo nibh arcu non elit. Maecenas maximus maximus justo, vitae bibendum orci semper eget.

Sed pulvinar leo eget nunc placerat, at cursus quam commodo. Vivamus nunc nisl, posuere id nunc ac, vehicula suscipit dui. Vivamus sodales urna felis, eget gravida nunc dignissim vel. Nam sed libero quis elit fermentum rhoncus. Vestibulum vitae libero et lectus ullamcorper congue eu in ex. In id ultricies enim. Integer volutpat nisi a lorem malesuada ultricies. Pellentesque ut commodo nulla. Nulla eu diam id lectus aliquet venenatis at quis ipsum.

Nulla ac mauris pharetra turpis interdum faucibus. Sed aliquam tellus ac finibus rutrum. Integer non vulputate sapien. Duis ullamcorper, felis non hendrerit interdum, ex lorem varius mauris, non consectetur eros tellus et ipsum. Nunc scelerisque turpis eu ullamcorper dictum. Cras ligula libero, pellentesque et pretium sed, tristique eget augue. Curabitur magna lectus, aliquam nec ullamcorper a, finibus pulvinar odio. Maecenas dictum, nisl et ornare porttitor, tortor metus molestie massa, ut porta nisi leo non velit. Maecenas quis rhoncus urna. Praesent a lacus vitae erat sagittis facilisis ac ac risus.

Quisque dapibus semper semper. Phasellus dignissim fringilla mollis. Duis volutpat lacus eget fringilla pretium. Phasellus ut orci non justo viverra aliquet sed at risus. Nulla eu tellus tincidunt, ultrices lectus vitae, tincidunt nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Nullam sem orci, fringilla sed nisi et, rhoncus sollicitudin odio. Nam sodales tortor ante, sit amet porttitor ipsum cursus id. Maecenas dapibus mauris nunc, ut placerat dui volutpat vitae. Donec viverra nulla nec libero iaculis mollis. Nullam viverra sit amet felis eu condimentum. In et tempor nibh. Phasellus interdum nibh ornare, scelerisque elit ac, fringilla arcu. Nullam rutrum vestibulum sollicitudin. Proin sem ante, tincidunt et nisi et, ultricies vulputate purus.

### 4. Player Skill Rating Model

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam fermentum sapien a felis mattis fermentum. Nunc laoreet eros in malesuada ultricies. Sed purus libero, blandit vitae convallis eu, lacinia ut justo. Curabitur est mauris, tincidunt in lacinia et, varius nec tortor. Nunc eu turpis feugiat, porta diam non, tincidunt lorem. Sed pulvinar, metus a volutpat malesuada, urna quam tempor nisi, nec cursus urna dolor vitae felis. Pellentesque posuere neque sapien, quis sagittis neque luctus quis. Quisque ullamcorper dignissim suscipit. Maecenas pharetra, nisi a pretium placerat, ante est posuere mauris, nec auctor mi diam in nisl. Pellentesque ligula magna, volutpat sed nunc vitae, auctor gravida arcu. Maecenas sit amet lacus justo.

In eu finibus purus. Etiam accumsan erat justo, eu efficitur elit faucibus non. In faucibus vitae nisl

dapibus ullamcorper. In fermentum leo eu purus rutrum suscipit. Pellentesque dictum efficitur purus, eu consequat metus pretium a. Vivamus feugiat urna ac accumsan suscipit. Praesent pretium sem risus, a congue felis placerat sed. Maecenas in nisi auctor, aliquet tortor ut, posuere justo. Cras lacinia bibendum elit, ut cursus turpis eleifend vitae. Nunc tempus libero quis suscipit scelerisque. Curabitur aliquet tempor facilisis. Suspendisse auctor, arcu non dictum luctus, sem tortor varius lorem, non commodo nibh arcu non elit. Maecenas maximus maximus justo, vitae bibendum orci semper eget.

Sed pulvinar leo eget nunc placerat, at cursus quam commodo. Vivamus nunc nisl, posuere id nunc ac, vehicula suscipit dui. Vivamus sodales urna felis, eget gravida nunc dignissim vel. Nam sed libero quis elit fermentum rhoncus. Vestibulum vitae libero et lectus ullamcorper congue eu in ex. In id ultricies enim. Integer volutpat nisi a lorem malesuada ultricies. Pellentesque ut commodo nulla. Nulla eu diam id lectus aliquet venenatis at quis ipsum.

Nulla ac mauris pharetra turpis interdum faucibus. Sed aliquam tellus ac finibus rutrum. Integer non vulputate sapien. Duis ullamcorper, felis non hendrerit interdum, ex lorem varius mauris, non consectetur eros tellus et ipsum. Nunc scelerisque turpis eu ullamcorper dictum. Cras ligula libero, pellentesque et pretium sed, tristique eget augue. Curabitur magna lectus, aliquam nec ullamcorper a, finibus pulvinar odio. Maecenas dictum, nisl et ornare porttitor, tortor metus molestie massa, ut porta nisi leo non velit. Maecenas quis rhoncus urna. Praesent a lacus vitae erat sagittis facilisis ac ac risus.

Quisque dapibus semper semper. Phasellus dignissim fringilla mollis. Duis volutpat lacus eget fringilla pretium. Phasellus ut orci non justo viverra aliquet sed at risus. Nulla eu tellus tincidunt, ultrices lectus vitae, tincidunt nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Nullam sem orci, fringilla sed nisi et, rhoncus sollicitudin odio. Nam sodales tortor ante, sit amet porttitor ipsum cursus id. Maecenas dapibus mauris nunc, ut placerat dui volutpat vitae. Donec viverra nulla nec libero iaculis mollis. Nullam viverra sit amet felis eu condimentum. In et tempor nibh. Phasellus interdum nibh ornare, scelerisque elit ac, fringilla arcu. Nullam rutrum vestibulum sollicitudin. Proin sem ante, tincidunt et nisi et, ultricies vulputate purus.

## 5. Future Work

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam fermentum sapien a felis mattis fermentum. Nunc laoreet eros in malesuada ultricies. Sed purus libero, blandit vitae convallis eu, lacinia ut justo. Curabitur est mauris, tincidunt in lacinia et, varius nec tortor. Nunc eu turpis feugiat, porta diam non, tincidunt lorem. Sed pulvinar, metus a volutpat malesuada, urna quam tempor nisi, nec cursus urna dolor vitae felis. Pellentesque posuere neque sapien, quis sagittis neque luctus quis. Quisque ullamcorper dignissim suscipit. Maecenas pharetra, nisi a pretium placerat, ante est posuere mauris, nec auctor mi diam in nisl. Pellentesque ligula magna, volutpat sed nunc vitae, auctor gravida arcu. Maecenas sit amet lacus justo.

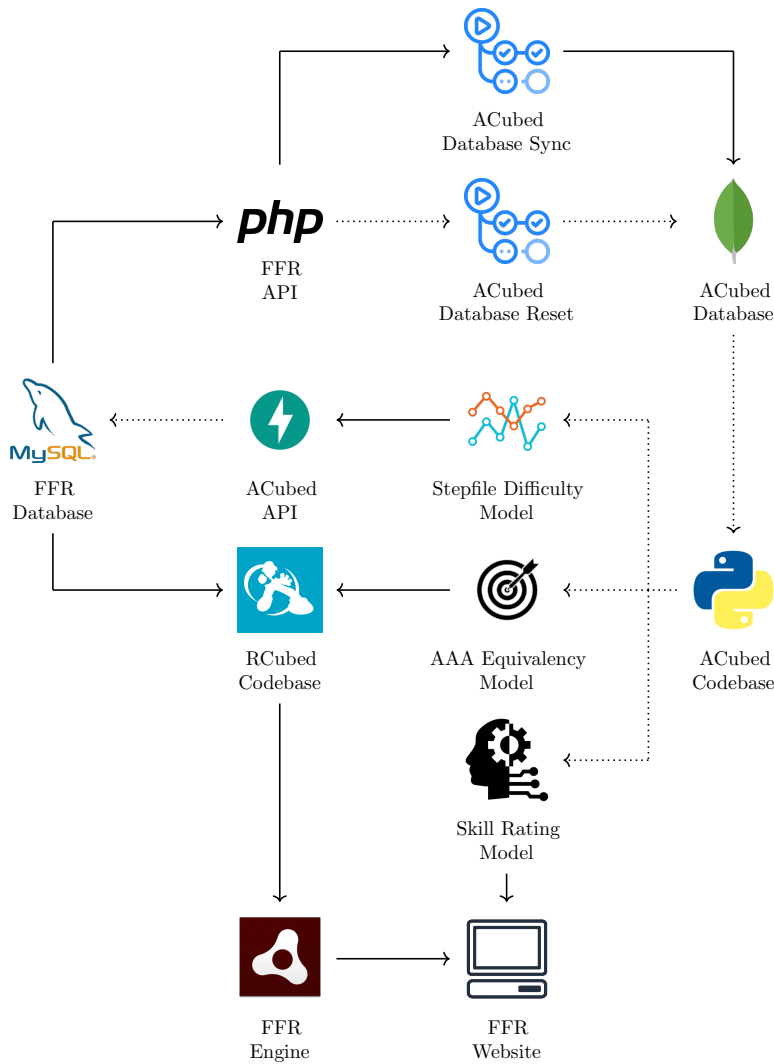
In eu finibus purus. Etiam accumsan erat justo, eu efficitur elit faucibus non. In faucibus vitae nisl dapibus ullamcorper. In fermentum leo eu purus rutrum suscipit. Pellentesque dictum efficitur purus, eu consequat metus pretium a. Vivamus feugiat urna ac accumsan suscipit. Praesent pretium sem risus, a congue felis placerat sed. Maecenas in nisi auctor, aliquet tortor ut, posuere justo. Cras lacinia bibendum elit, ut cursus turpis eleifend vitae. Nunc tempus libero quis suscipit scelerisque. Curabitur aliquet tempor facilisis. Suspendisse auctor, arcu non dictum luctus, sem tortor varius lorem, non commodo nibh arcu non elit. Maecenas maximus maximus justo, vitae bibendum orci semper eget.

Sed pulvinar leo eget nunc placerat, at cursus quam commodo. Vivamus nunc nisl, posuere id nunc ac, vehicula suscipit dui. Vivamus sodales urna felis, eget gravida nunc dignissim vel. Nam sed libero quis elit fermentum rhoncus. Vestibulum vitae libero et lectus ullamcorper congue eu in ex. In id ultricies enim. Integer volutpat nisi a lorem malesuada ultricies. Pellentesque ut commodo nulla. Nulla eu diam id lectus aliquet venenatis at quis ipsum.

Nulla ac mauris pharetra turpis interdum faucibus. Sed aliquam tellus ac finibus rutrum. Integer non vulputate sapien. Duis ullamcorper, felis non hendrerit interdum, ex lorem varius mauris, non consectetur eros tellus et ipsum. Nunc scelerisque turpis eu ullamcorper dictum. Cras ligula libero, pellentesque et pretium sed, tristique eget augue. Curabitur magna lectus, aliquam nec ullamcorper a, finibus pulvinar odio. Maecenas dictum, nisl et ornare porttitor, tortor metus molestie massa, ut porta nisi leo non velit. Maecenas quis rhoncus urna. Praesent a lacus vitae erat sagittis facilisis ac ac risus.

Quisque dapibus semper semper. Phasellus dignissim fringilla mollis. Duis volutpat lacus eget fringilla pretium. Phasellus ut orci non justo viverra aliquet sed at risus. Nulla eu tellus tincidunt, ultrices lectus vitae, tincidunt nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Nullam sem orci, fringilla sed nisi et, rhoncus sollicitudin odio. Nam sodales tortor ante, sit amet porttitor ipsum cursus id. Maecenas dapibus mauris nunc, ut placerat dui volutpat vitae. Donec viverra nulla nec libero iaculis mollis. Nullam viverra sit amet felis eu condimentum. In et tempor nibh. Phasellus interdum nibh ornare, scelerisque elit ac, fringilla arcu. Nullam rutrum vestibulum sollicitudin. Proin sem ante, tincidunt et nisi et, ultricies vulputate purus.

## 6. Technical Implementation



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam fermentum sapien a felis mattis fermentum. Nunc laoreet eros in malesuada ultricies. Sed purus libero, blandit vitae convallis eu, lacinia ut justo. Curabitur est mauris, tincidunt in lacinia et, varius nec tortor. Nunc eu turpis feugiat, porta diam non, tincidunt lorem. Sed pulvinar, metus a volutpat malesuada, urna quam tempor nisi, nec cursus urna dolor vitae felis. Pellentesque posuere neque sapien, quis sagittis neque luctus quis. Quisque ullamcorper dignissim suscipit. Maecenas pharetra, nisi a pretium placerat, ante est posuere mauris, nec auctor mi diam in nisl. Pellentesque ligula magna, volutpat sed nunc vitae, auctor gravida arcu. Maecenas sit amet lacus justo.

In eu finibus purus. Etiam accumsan erat justo, eu efficitur elit faucibus non. In faucibus vitae nisl dapibus ullamcorper. In fermentum leo eu purus rutrum suscipit. Pellentesque dictum efficitur purus, eu consequat metus pretium a. Vivamus feugiat urna ac accumsan suscipit. Praesent pretium sem risus, a congue felis placerat sed. Maecenas in nisi auctor, aliquet tortor ut, posuere justo. Cras lacinia bibendum elit, ut cursus turpis eleifend vitae. Nunc tempus libero quis suscipit scelerisque. Curabitur aliquet tempor facilisis. Suspendisse auctor, arcu non dictum luctus, sem tortor varius lorem, non

commodo nibh arcu non elit. Maecenas maximus maximus justo, vitae bibendum orci semper eget.

Sed pulvinar leo eget nunc placerat, at cursus quam commodo. Vivamus nunc nisl, posuere id nunc ac, vehicula suscipit dui. Vivamus sodales urna felis, eget gravida nunc dignissim vel. Nam sed libero quis elit fermentum rhoncus. Vestibulum vitae libero et lectus ullamcorper congue eu in ex. In id ultricies enim. Integer volutpat nisi a lorem malesuada ultricies. Pellentesque ut commodo nulla. Nulla eu diam id lectus aliquet venenatis at quis ipsum.

Nulla ac mauris pharetra turpis interdum faucibus. Sed aliquam tellus ac finibus rutrum. Integer non vulputate sapien. Duis ullamcorper, felis non hendrerit interdum, ex lorem varius mauris, non consectetur eros tellus et ipsum. Nunc scelerisque turpis eu ullamcorper dictum. Cras ligula libero, pellentesque et pretium sed, tristique eget augue. Curabitur magna lectus, aliquam nec ullamcorper a, finibus pulvinar odio. Maecenas dictum, nisl et ornare porttitor, tortor metus molestie massa, ut porta nisi leo non velit. Maecenas quis rhoncus urna. Praesent a lacus vitae erat sagittis facilisis ac ac risus.

Quisque dapibus semper semper. Phasellus dignissim fringilla mollis. Duis volutpat lacus eget fringilla pretium. Phasellus ut orci non justo viverra aliquet sed at risus. Nulla eu tellus tincidunt, ultrices lectus vitae, tincidunt nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Nullam sem orci, fringilla sed nisi et, rhoncus sollicitudin odio. Nam sodales tortor ante, sit amet porttitor ipsum cursus id. Maecenas dapibus mauris nunc, ut placerat dui volutpat vitae. Donec viverra nulla nec libero iaculis mollis. Nullam viverra sit amet felis eu condimentum. In et tempor nibh. Phasellus interdum nibh ornare, scelerisque elit ac, fringilla arcu. Nullam rutrum vestibulum sollicitudin. Proin sem ante, tincidunt et nisi et, ultricies vulputate purus.

## 7. Conclusion and Acknowledgements

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam fermentum sapien a felis mattis fermentum. Nunc laoreet eros in malesuada ultricies. Sed purus libero, blandit vitae convallis eu, lacinia ut justo. Curabitur est mauris, tincidunt in lacinia et, varius nec tortor. Nunc eu turpis feugiat, porta diam non, tincidunt lorem. Sed pulvinar, metus a volutpat malesuada, urna quam tempor nisi, nec cursus urna dolor vitae felis. Pellentesque posuere neque sapien, quis sagittis neque luctus quis. Quisque ullamcorper dignissim suscipit. Maecenas pharetra, nisi a pretium placerat, ante est posuere mauris, nec auctor mi diam in nisl. Pellentesque ligula magna, volutpat sed nunc vitae, auctor gravida arcu. Maecenas sit amet lacus justo.

In eu finibus purus. Etiam accumsan erat justo, eu efficitur elit faucibus non. In faucibus vitae nisl dapibus ullamcorper. In fermentum leo eu purus rutrum suscipit. Pellentesque dictum efficitur purus, eu consequat metus pretium a. Vivamus feugiat urna ac accumsan suscipit. Praesent pretium sem risus, a congue felis placerat sed. Maecenas in nisi auctor, aliquet tortor ut, posuere justo. Cras lacinia bibendum elit, ut cursus turpis eleifend vitae. Nunc tempus libero quis suscipit scelerisque. Curabitur aliquet tempor facilisis. Suspendisse auctor, arcu non dictum luctus, sem tortor varius lorem, non commodo nibh arcu non elit. Maecenas maximus maximus justo, vitae bibendum orci semper eget.

Sed pulvinar leo eget nunc placerat, at cursus quam commodo. Vivamus nunc nisl, posuere id nunc ac, vehicula suscipit dui. Vivamus sodales urna felis, eget gravida nunc dignissim vel. Nam sed libero quis elit fermentum rhoncus. Vestibulum vitae libero et lectus ullamcorper congue eu in ex. In id ultricies enim. Integer volutpat nisi a lorem malesuada ultricies. Pellentesque ut commodo nulla. Nulla eu diam id lectus aliquet venenatis at quis ipsum.

Nulla ac mauris pharetra turpis interdum faucibus. Sed aliquam tellus ac finibus rutrum. Integer non vulputate sapien. Duis ullamcorper, felis non hendrerit interdum, ex lorem varius mauris, non consectetur eros tellus et ipsum. Nunc scelerisque turpis eu ullamcorper dictum. Cras ligula libero, pellentesque et pretium sed, tristique eget augue. Curabitur magna lectus, aliquam nec ullamcorper a,



finibus pulvinar odio. Maecenas dictum, nisl et ornare porttitor, tortor metus molestie massa, ut porta nisi leo non velit. Maecenas quis rhoncus urna. Praesent a lacus vitae erat sagittis facilisis ac ac risus.

Quisque dapibus semper semper. Phasellus dignissim fringilla mollis. Duis volutpat lacus eget fringilla pretium. Phasellus ut orci non justo viverra aliquet sed at risus. Nulla eu tellus tincidunt, ultrices lectus vitae, tincidunt nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Nullam sem orci, fringilla sed nisi et, rhoncus sollicitudin odio. Nam sodales tortor ante, sit amet porttitor ipsum cursus id. Maecenas dapibus mauris nunc, ut placerat dui volutpat vitae. Donec viverra nulla nec libero iaculis mollis. Nullam viverra sit amet felis eu condimentum. In et tempor nibh. Phasellus interdum nibh ornare, scelerisque elit ac, fringilla arcu. Nullam rutrum vestibulum sollicitudin. Proin sem ante, tincidunt et nisi et, ultricies vulputate purus.

## A. Appendix

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam fermentum sapien a felis mattis fermentum. Nunc laoreet eros in malesuada ultricies. Sed purus libero, blandit vitae convallis eu, lacinia ut justo. Curabitur est mauris, tincidunt in lacinia et, varius nec tortor. Nunc eu turpis feugiat, porta diam non, tincidunt lorem. Sed pulvinar, metus a volutpat malesuada, urna quam tempor nisi, nec cursus urna dolor vitae felis. Pellentesque posuere neque sapien, quis sagittis neque luctus quis. Quisque ullamcorper dignissim suscipit. Maecenas pharetra, nisi a pretium placerat, ante est posuere mauris, nec auctor mi diam in nisl. Pellentesque ligula magna, volutpat sed nunc vitae, auctor gravida arcu. Maecenas sit amet lacus justo.

In eu finibus purus. Etiam accumsan erat justo, eu efficitur elit faucibus non. In faucibus vitae nisl dapibus ullamcorper. In fermentum leo eu purus rutrum suscipit. Pellentesque dictum efficitur purus, eu consequat metus pretium a. Vivamus feugiat urna ac accumsan suscipit. Praesent pretium sem risus, a congue felis placerat sed. Maecenas in nisi auctor, aliquet tortor ut, posuere justo. Cras lacinia bibendum elit, ut cursus turpis eleifend vitae. Nunc tempus libero quis suscipit scelerisque. Curabitur aliquet tempor facilisis. Suspendisse auctor, arcu non dictum luctus, sem tortor varius lorem, non commodo nibh arcu non elit. Maecenas maximus maximus justo, vitae bibendum orci semper eget.

Sed pulvinar leo eget nunc placerat, at cursus quam commodo. Vivamus nunc nisl, posuere id nunc ac, vehicula suscipit dui. Vivamus sodales urna felis, eget gravida nunc dignissim vel. Nam sed libero quis elit fermentum rhoncus. Vestibulum vitae libero et lectus ullamcorper congue eu in ex. In id ultricies enim. Integer volutpat nisi a lorem malesuada ultricies. Pellentesque ut commodo nulla. Nulla eu diam id lectus aliquet venenatis at quis ipsum.

Nulla ac mauris pharetra turpis interdum faucibus. Sed aliquam tellus ac finibus rutrum. Integer non vulputate sapien. Duis ullamcorper, felis non hendrerit interdum, ex lorem varius mauris, non consectetur eros tellus et ipsum. Nunc scelerisque turpis eu ullamcorper dictum. Cras ligula libero, pellentesque et pretium sed, tristique eget augue. Curabitur magna lectus, aliquam nec ullamcorper a, finibus pulvinar odio. Maecenas dictum, nisl et ornare porttitor, tortor metus molestie massa, ut porta nisi leo non velit. Maecenas quis rhoncus urna. Praesent a lacus vitae erat sagittis facilisis ac ac risus.

Quisque dapibus semper semper. Phasellus dignissim fringilla mollis. Duis volutpat lacus eget fringilla pretium. Phasellus ut orci non justo viverra aliquet sed at risus. Nulla eu tellus tincidunt, ultrices lectus vitae, tincidunt nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Nullam sem orci, fringilla sed nisi et, rhoncus sollicitudin odio. Nam sodales tortor ante, sit amet porttitor ipsum cursus id. Maecenas dapibus mauris nunc, ut placerat dui volutpat vitae. Donec viverra nulla nec libero iaculis mollis. Nullam viverra sit amet felis eu condimentum. In et tempor nibh. Phasellus interdum nibh ornare, scelerisque elit ac, fringilla arcu. Nullam rutrum vestibulum sollicitudin. Proin sem ante, tincidunt et nisi et, ultricies vulputate purus.

## References

- [nAS]    nois-or-e, One Winged Angel, and Stavie33. *General Difficulty Discussions & Consultant Insight*. <https://www.flashflashrevolution.com/vbz/showthread.php?t=124260>. Created: 2012-06-08.
- [Pra]    PrawkSkunk. *NEW Skill-Based Ranking System & Leaderboards*. <https://www.flashflashrevolution.com/ffr/new-skill-based-ranking-system-leaderboards/>. Created: 2015-02-16.
- [PXW]    PrawnSkunk, XJ-9, and WirryWoo. *FFR Season 1 Leaderboards*. <https://www.flashflashrevolution.com/ffr/season-1/>. Created: 2021-10-01.
- [Tea]    The FFR Development Team. *RCubed Goes Open Source!* <https://www.flashflashrevolution.com/ffr/r3-goes-open-source/>. Created: 2020-04-27.

- [Trua] Trumpet63. *I made an auto-difficulty model*. <https://www.flashflashrevolution.com/vbz/showthread.php?t=154255>. Created: 2023-02-16.
- [Trub] Trumpet63. *New Skill Analysis Metric (Trial Phase)*. <https://www.flashflashrevolution.com/vbz/showthread.php?p=4209451>. Created: 2014-09-30.