

# R6 Seige Report

Makenzy Gilbert

2023-10-26

## Rainbow Six Siege Recoil Analysis

### Introduction

Rainbow Six Siege is a first-person shooter video game with two teams of five players each. One team acts as the “attackers,” striving to complete objectives such as rescuing hostages or defusing bombs, while the other team serves as the “defenders,” aiming to thwart their efforts. Before each match, players select their “operator,” each of whom possesses a unique gadget or ability that significantly impacts the game’s strategy.

During the match, both teams can ban one attacking and one defending operator, preventing the opposing team from selecting those operators. This operator selection process adds an extra layer of strategy to the game. The rounds are fast-paced, requiring meticulous planning and quick reflexes.

Operators wield firearms with specific recoil patterns, mirroring their real-world counterparts. Some operators share the same primary weapons, resulting in shared recoil patterns. These patterns dictate how a weapon behaves when fired continuously and influence aim deviation. Mastering these recoil patterns enables players to maintain accuracy during intense firefights.

Rainbow Six Siege demands players’ full attention, making sound an integral part of the game. With over 800 unique audio cues to discern, sound cues play a crucial role in gameplay. Additionally, numerous environmental objects offer tactical advantages, requiring players to monitor various angles for opponents. Players can also customize their weapon sights or scopes, further affecting their aiming mechanics. Memorizing maps, room layouts, and operator abilities adds another layer of strategic depth to the game.

Incorporating principles from cognitive load theory, game developers recognize the importance of working memory limitations in instructional design. This theory aligns with Miller’s Law, which suggests that people can hold around 7 (plus or minus 2) items in their working memory. Fitt’s Law, which states that the time to acquire a target depends on its size and distance, also informs game design decisions.

In the context of Rainbow Six Siege, weapons with wider recoil spreads demand more attention and control, aligning with the concept of allocating attentional resources in-game. Game development studios employ seasoned UX/UI researchers to enhance game play by considering cognitive limitations and abilities, making Rainbow Six Siege a strategically engaging and immersive gaming experience.

Research Question: Does higher levels of cognitive strain due to larger spread in recoil patterns (of weapons for operators), influence or correlate with Pick vs. Ban Rates of operators?

## Setup

*# Dependencies*

```
library(tidyverse)
```

```
## — Attaching core tidyverse packages ————— tidyverse  
2.0.0 —
```

```
## ✓ dplyr      1.1.3      ✓ readr      2.1.4
```

```
## ✓ forcats   1.0.0      ✓ stringr   1.5.0
```

```
## ✓ ggplot2   3.4.4      ✓ tibble    3.2.1
```

```
## ✓ lubridate 1.9.3      ✓ tidyr     1.3.0
```

```
## ✓ purrr     1.0.2
```

```
## — Conflicts —————
```

```
tidyverse_conflicts() —
```

```
## ✗ dplyr::filter() masks stats::filter()
```

```
## ✗ dplyr::lag()     masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all  
conflicts to become errors
```

```
library(ggrepel)
```

```
library(skimr)
```

```
library(rvest)
```

```
##
```

```
## Attaching package: 'rvest'
```

```
##
```

```
## The following object is masked from 'package:readr':
```

```
##
```

```
##   guess_encoding
```

```
library(google sheets4)
```

```
library(viridis)
```

```
## Loading required package: viridisLite
```

```
library(hexbin)
```

```
gs4_auth()
```

```
## ! Using an auto-discovered, cached token.
```

```
## To suppress this message, modify your code or options to clearly consent  
to
```

```
## the use of a cached token.
```

```
## See gargle's "Non-interactive auth" vignette for more details:
```

```
## <https://gargle.r-lib.org/articles/non-interactive-auth.html>
```

```
## i The googlesheets4 package is using a cached token for
```

```
'makenzy101@gmail.com'.
```

```
# googlesheets4::gs4_auth(force = TRUE)
```

## Methods

### Data Source #1

Rvesting the data set:

This data set consists of a table of all weapons used in the game and provides statistics; such as, type of gun, firing rate, and the operator that uses it. The original creators of the game and the publishing company, Ubisoft, are overall the generators for the data I am using. The first data source, IGN, provides information about weapons and equipment in Rainbow Six Siege. IGN is a popular gaming and entertainment website, and they compiled this data to provide a resource for gamers interested in the game's details.

```
url <- "https://www.ign.com/wikis/rainbow-six-siege/Weapons_and_Equipment"
page <- read_html(url)
table_data <- page %>%
  html_table()
rawweapons_table <- table_data[[1]]
```

### Check NA's and Packaging

```
skim(table_data)
```

*Data summary*

Name	table_data
Number of rows	100
Number of columns	11

---

Column type frequency:

character	10
numeric	1

---

Group variables	None
-----------------	------

**Variable type: character**

skim_variable	n_missin g	complete_rat e	mi n	ma x	empt y	n_uniqu e	whitespac e
Name	0	1	2	17	0	100	0
Type	0	1	3	18	0	8	0
Operator	0	1	2	29	0	69	0

skim_variable	n_missin g	complete_rat e	mi n	ma x	empt y	n_uniqu e	whitespac e
Team	0	1	8	16	0	3	0
CTU	0	1	3	23	0	29	0
Damage..Suppresse d..	0	1	2	7	0	50	0
ROF	0	1	3	9	0	33	0
Suppressor	0	1	1	2	0	3	0
ACOG	0	1	1	2	0	3	0
Var.11	0	1	9	19	0	7	0

### Variable type: numeric

skim_variab le	n_missin g	complete_ra te	mea n	sd	p 0	p2 5	p5 0	p7 5	p10 0	hist
Magazine	0	1	25.2 8	22.1 7	2	10	20	30	150	█ _ _ _ —

There is one column missing a name

Remove/Replace NA in column names:

```
View(rawweapons_table)
colnames(rawweapons_table)

## [1] "Name"           "Type"           "Operator"
## [4] "Team"           "CTU"            "Damage (Suppressed*)"
## [7] "ROF"            "Magazine"       "Suppressor"
## [10] "ACOG"           ""

# [1] "Name"           "Type"           "Operator"
# [4] "Team"           "CTU"            "Damage (Suppressed*)"
# [7] "ROF"            "Magazine"       "Suppressor"
# [10] "ACOG"           ""

weapons_table <- rawweapons_table

colnames(weapons_table)[11] <- "Range"
```

Next I need to rename “Name” column to be more descriptive:

```
weapons_table1 <- weapons_table %>%
  rename(Weapon = Name)
```

The next packaging issue is the concatenated “overlapping” of operators per each weapon. I need to split these.

```
weapons_table1 <- weapons_table1 %>%
  mutate(Operator = ifelse(Operator == "Jäger", "Jager", Operator))
```

```

# Splitting the operator column based on uppercase letters
# Split the Operator column and unnest the result
weapons_table1_split <- weapons_table1 %>%
  # Use str_extract_all to find all uppercase followed by lowercase strings
  # (operators' names)
  # or all uppercase strings with 2 or more characters
  mutate(Operator = str_extract_all(Operator, "([A-Z][a-z]+|[A-Z]{2,})")) %>%
  # Unnest the list column to separate rows
  unnest(Operator)

```

## Data Source #2

The second main data source, Gameriv, offers pick and ban rates for operators in Rainbow Six Siege. Gameriv is a gaming-focused website, and they collected this data to inform the gaming community about the popularity and ban rates of operators in the game. (Graphs were collected straight from Ubisoft's seasonal data roundup.) The exact data was collected from seasonal statistics of gameplay from year 7 early few seasons 1.2-3. released March of 2022 and data from the designer's notes officially was published on Ubisoft's website 4/5/2022. The Gameriv review article was published September 12, 2022.

From Source: "Credit: All the graphs below have been collected from Ubisoft and contain data recorded from PCRanked Platinum and above." This is helpful to note that the data comes from players using only a personal computer in high level ranked games, as opposed to console players being included.

These ranks are also important to note. Overall, the player base that the data stems from is ranked by these ranks: (worst to best).

- Copper V, IV, III, II, I
- Bronze V, IV, III, II, I
- Silver V, IV, III, II, I
- Gold V, IV, III, II, I
- Platinum V, IV, III, II, I
- Emerald V, IV, III, II, I
- Diamond V, IV, III, II, I
- Champions

My data comes from the upper echelon of skilled players. It is the top half of all ranked players that provide this data. It is a set of extremely knowledgeable and experienced players. This is important due to the fact that these sets of players attention is extremely focused in-game. So, to investigate underlying tendencies to minimize attentional demand or "cognitive load", I used this set of pick rate data.

High pick, average pick, and low pick rates have been determined by a metric as follows from Ubisoft: Percentage per x amount of won games:

Attackers: High Pick Rate (Presence > 40%) Average Pick Rate (15% < Presence < 40%)  
 Low Pick Rate (Presence < 15%)

Defenders: High Pick Rate (Presence > 30%) Average Pick Rate (18% < Presence < 30%)  
Low Pick Rate (Presence < 10%)

### Add pick rate column

```
high_pick <- c("Finka", "Iana", "Sledge", "Thatcher", "Twitch", "Alibi",  
"Jager", "Melusi", "Mute")  
average_pick <- c("Valkyrie", "Thunderbird", "Smoke", "Bandit", "Kaid",  
"Wamai", "Mozzie", "Azami", "Lesion", "Kapkan", "Aruni", "Oryx", "Jackal",  
"Ace", "Thermite", "Zofia", "Maverick", "Hibana", "Ash", "Nomad")  
low_pick <- c("Buck", "Osa", "Nokk", "Flores", "Amaru", "IQ", "Zero",  
"Montagne", "Dokkaebi", "Blackbeard", "Lion", "Blitz", "Kali", "Capitao",  
"Fuze", "Ying", "Gridlock", "Glaz", "Sens", "Vigil", "Doc", "Ela", "Mira",  
"Goyo", "Frost", "Thorn", "Maestro", "Pulse", "Rook", "Castle", "Clash",  
"Warden", "Caveira", "Echo", "Tachanka")
```

```
weaponstable_long <- weapons_table1_split %>%  
  mutate(PickRate = case_when(  
    Operator %in% high_pick ~ "High Pick Rate",  
    Operator %in% average_pick ~ "Average Pick Rate",  
    Operator %in% low_pick ~ "Low Pick Rate",  
    TRUE ~ NA_character_  
  ))
```

*# View the modified table with the new PickRate column*  
**summary**(weaponstable\_long)

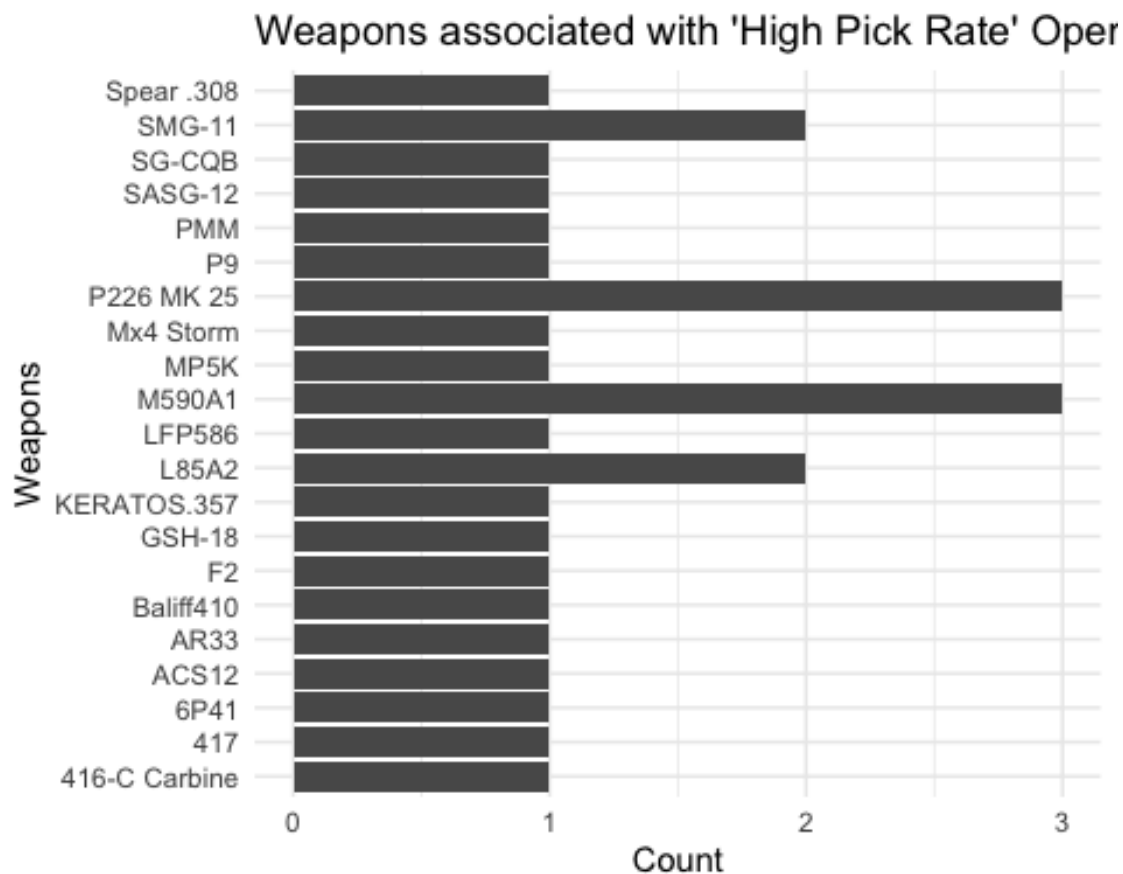
```
##      Weapon              Type      Operator      Team  
## Length:168      Length:168      Length:168      Length:168  
## Class :character Class :character Class :character Class :character  
## Mode  :character Mode  :character Mode  :character Mode  :character  
##  
##  
##  
##      CTU      Damage (Suppressed*)      ROF      Magazine  
## Length:168      Length:168      Length:168      Min.   : 2.00  
## Class :character Class :character      Class :character      1st Qu.: 8.00  
## Mode  :character Mode  :character      Mode  :character      Median :16.00  
##                                     Mean   :21.23  
##                                     3rd Qu.:30.00  
##                                     Max.   :150.00  
##  
##      Suppressor      ACOG      Range      PickRate  
## Length:168      Length:168      Length:168      Length:168  
## Class :character Class :character      Class :character      Class :character  
## Mode  :character Mode  :character      Mode  :character      Mode  :character  
##  
##  
##
```

*# Filter for High Pick Rate weapons*

```
high_pick_rate_weapons <- weaponstable_long %>% filter(`PickRate` == 'High
```

```
Pick Rate')

# Create the ggplot
ggplot(high_pick_rate_weapons, aes(x = Weapon)) +
  geom_bar() +
  coord_flip() + # Flipping the axis for better visualization
  labs(title = "Weapons associated with 'High Pick Rate' Operators",
       x = "Weapons",
       y = "Count") +
  theme_minimal()
```



This visual shows all weapons that are used by Highly Picked operators. The count gives a measure of how many high-pick operators use the weapon (Indicates multiple high pick operators use the weapon).

### Data Source #3

#### Recoil statistics

These statistics come from previous analysis here:

<https://medium.com/@DogtorFlashbank/which-attacker-guns-in-siege-are-the-best-9dac3e7f7688> This source provides a score for each weapons recoil statistic. The purpose of this data was to provide insights over all variables effecting what makes an operator

strong or weak in game play. They found this score by focusing on automatic-firing weapons, and looking at the 4th successive bullets distance from the center at start of firing, they measured how high (y) and how far to the side (x) it reached, in pixels.

Vertical and Horizontal scores were calculated and summed for each weapon and recorded on a scale of the smallest recoil gets one point, the bigger gets closer to zero points and into the negative points for the worst. Horizontal movement gets a weight of 0.9 for and 0.6 for vertical (1.5x both weights then averaged) to reach the overall score for recoil. This means in total a gun can get about point total for recoil.

It is important to note that the original analyst used each weapon with the same "equipments" to control consistency.

I want to extract column "L" containing the recoil scores, as well as "B" for the associated weapon. There are two separate sheets one is attacker and one is for defenders so I will repeat the extraction for both sheets.

```
score<- "1BxIHO_VHKXuNYypY6NRihik98HJN1b9d0_nZCE_1-ts"
score_arecoil <-googlesheets4::read_sheet(score, range = "attacker guns!L:L")

## ✓ Reading from "Copy of SCOAR: the best guns in Siege, by Dogtor Flashbank".

## ✓ Range ''attacker guns'!L:L'.

score_aweapon <- read_sheet(score, range = "attacker guns!B:B")

## ✓ Reading from "Copy of SCOAR: the best guns in Siege, by Dogtor Flashbank".

## ✓ Range ''attacker guns'!B:B'.

# Merge the two columns by row number
atrecoil_data <- cbind(score_aweapon, score_arecoil)

# View the merged data frame
print(atrecoil_data)

score_drecoil <-googlesheets4::read_sheet(score, range = "defender guns!L:L")

## ✓ Reading from "Copy of SCOAR: the best guns in Siege, by Dogtor Flashbank".

## ✓ Range ''defender guns'!L:L'.

score_dweapon <- read_sheet(score, range = "defender guns!B:B")

## ✓ Reading from "Copy of SCOAR: the best guns in Siege, by Dogtor Flashbank".

## ✓ Range ''defender guns'!B:B'.
```



```

# Merge the two columns by row number
defrecoil_data <- cbind(score_dweapon, score_drecoil)

# View the merged data frame
print(defrecoil_data)

combo_recoil <- rbind(defrecoil_data, atrecoil_data)
combo_recoilunique<- unique(combo_recoil)
## interesting 71 obs... still some missing,

```

After acquiring these scores, I need to assign recoil statistics to each of the weapons in my dataset. But upon inspection there are missing guns besides the handguns. Is that fit for my research still?

There were slight discrepancies in spelling for each weapon, but besides that, some of the weapon data was not fit for my research question.

Handguns and machine pistols in this dataset are used as “Secondary” weapons in the game. This is not what I need to look at because that is not the main weapon for the operators. Marksman Rifles have relatively little to NO kick at all between shots due to the fact that they function as single-fire weapons. So in this case I want to filter them out because they do not drift from target because they’re not automatic weapons.

Shotguns have massive variability again in the functionality. These weapons fire single “rounds” that have a spread of bullets in a blob. Not a weapon that calls for countering recoil in the same manner as auto-fire weapons, so we will filter these out as well.

```

# Filtering the table
filtered_table <- weaponstable_long %>%
  filter(
    Weapon != "AR-15.50",
    Type != "Handgun",
    Type != "Marksman Rifle",
    Type != "Shotgun",
    !(Type == "Machine Pistol" & Weapon != "SPSMG9")
  )

# Get the unique weapons from the Weapon column after filtering
unique_weapons <- unique(filtered_table$Weapon)

# Count the number of unique weapons
number_of_unique_weapons <- length(unique_weapons)

# Print the result
print(number_of_unique_weapons)

## [1] 50

```

Now that the dimensions match from the combined recoil scores data frame and the weapons in the original data frame, and the “unfit” weapons are removed I merged the two data frames

```
fulldf<- merge(filtered_table, combo_recoilunique, by = "Weapon", all.x = TRUE)

#print(fulldf)
```

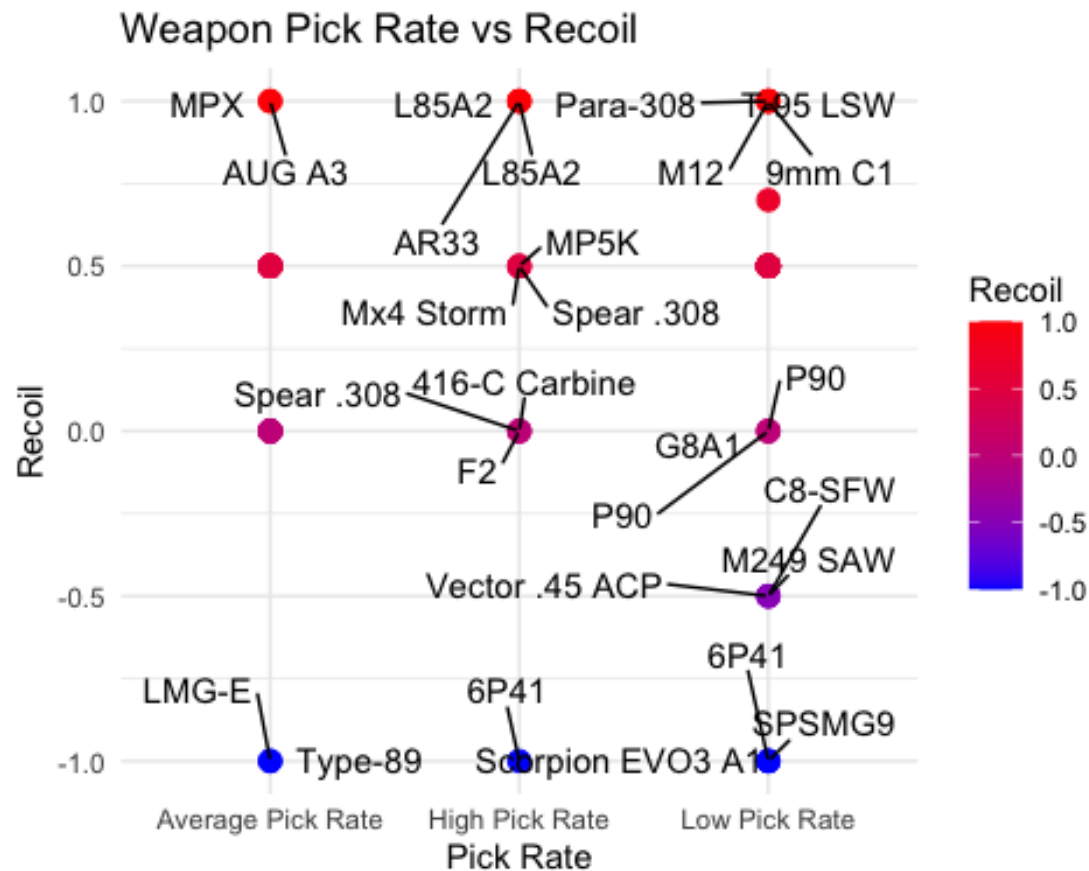
## Results

I want a clear visual that conveys how many weapons are associated with high pick rate operators as compared to the corresponding recoil score of the weapon. I started by trying to look at frequency:

```
# Discretize the Recoil column
fulldf$RecoilRange <- cut(fulldf$Recoil,
                          breaks=c(-Inf, -0.5, 0, 0.5, Inf),
                          labels=c("Very High Recoil", "High Recoil", "Medium Recoil", "Low Recoil"))

ggplot(fulldf, aes(x = PickRate, y = Recoil, label = Weapon)) +
  geom_point(aes(color = Recoil), size = 3) +
  geom_text_repel(aes(label = Weapon), box.padding = 0.5) +
  labs(title = "Weapon Pick Rate vs Recoil",
       x = "Pick Rate",
       y = "Recoil") +
  theme_minimal() +
  scale_color_gradient(low = "blue", high = "red")

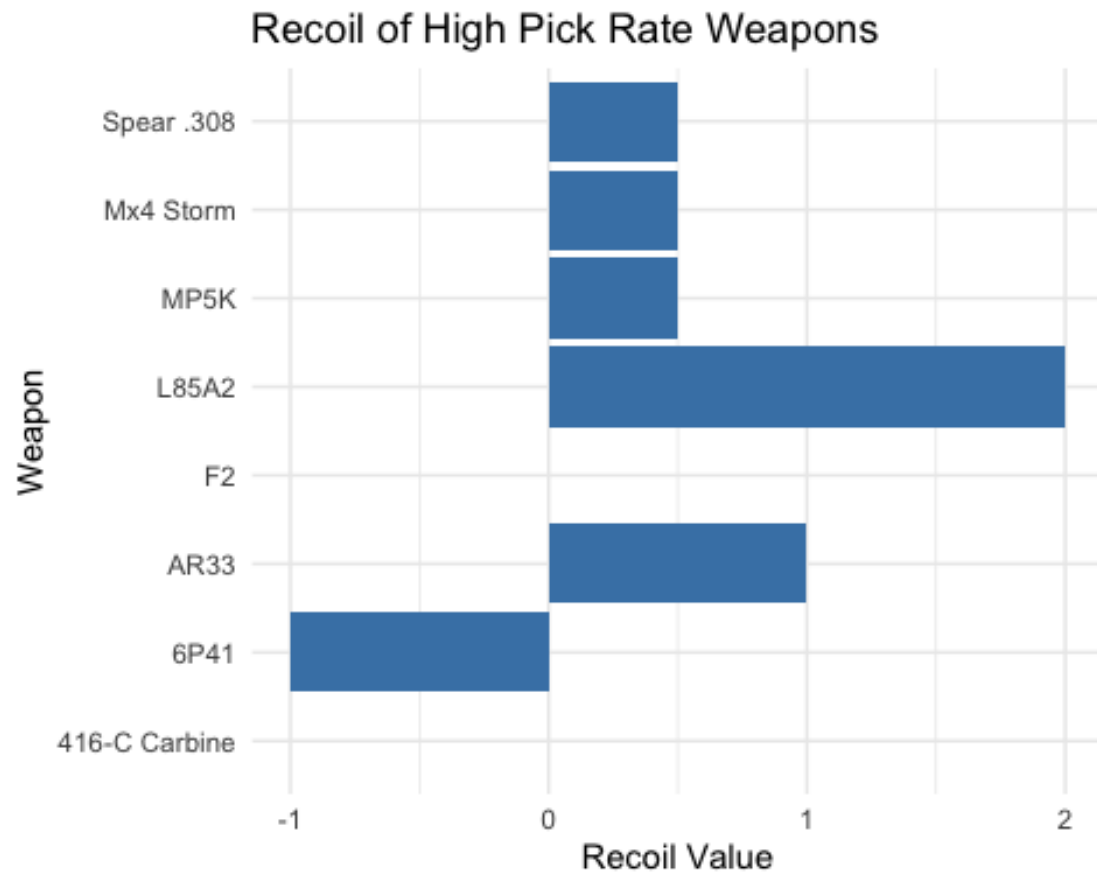
## Warning: ggrepel: 30 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



The temperature scale on the right as we will call it, is slightly confusing. “High recoil” interpreted in the data is actually optimal, but the phrase alone suggests “more” or “higher” recoil. So this is not the best representation of results

```
high_pick_df <- subset(full_df, PickRate == "High Pick Rate")

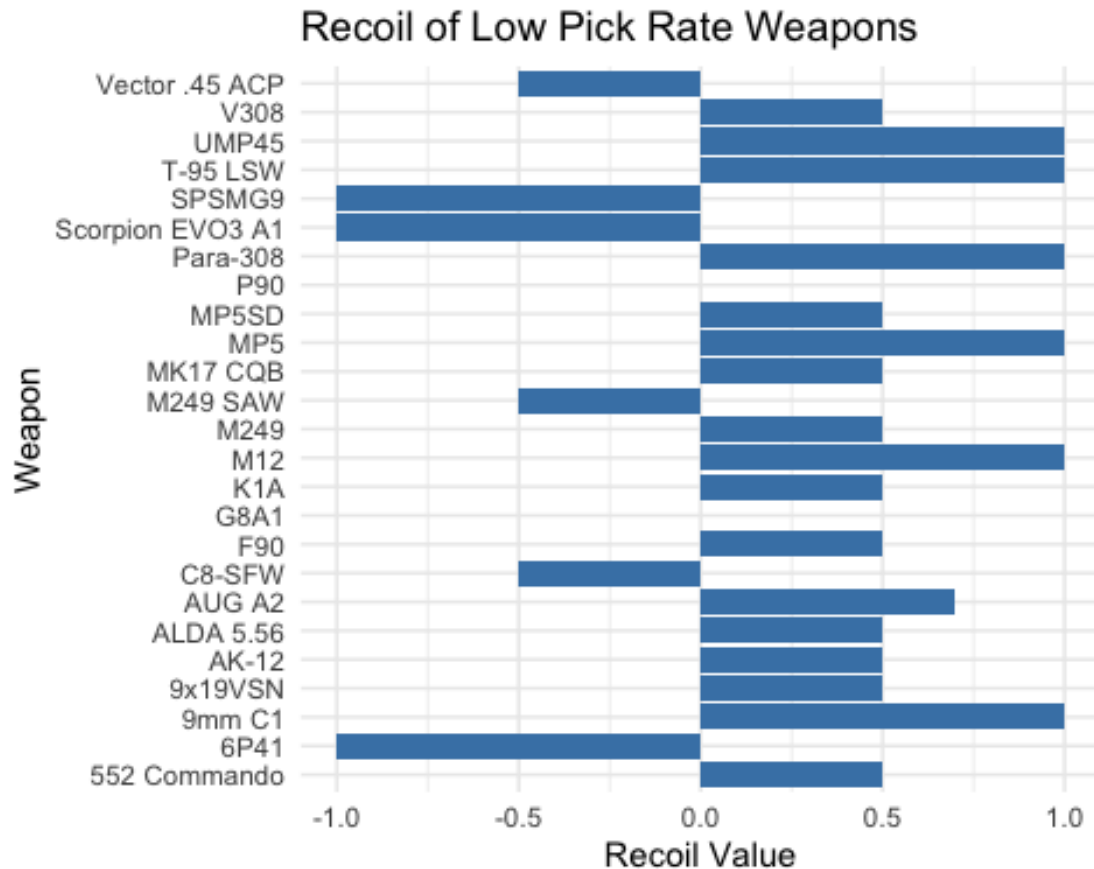
ggplot(high_pick_df, aes(x = Weapon, y = Recoil)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  coord_flip() + # Makes the plot horizontal for better readability with
  weapon names
  labs(title = "Recoil of High Pick Rate Weapons",
       x = "Weapon",
       y = "Recoil Value") +
  theme_minimal()
```



This is a measure of spread that informatively shows a mostly - positive distribution with high-pick weapons recoil scores being positive.

```
low_pick_df <- subset(full_df, PickRate == "Low Pick Rate")

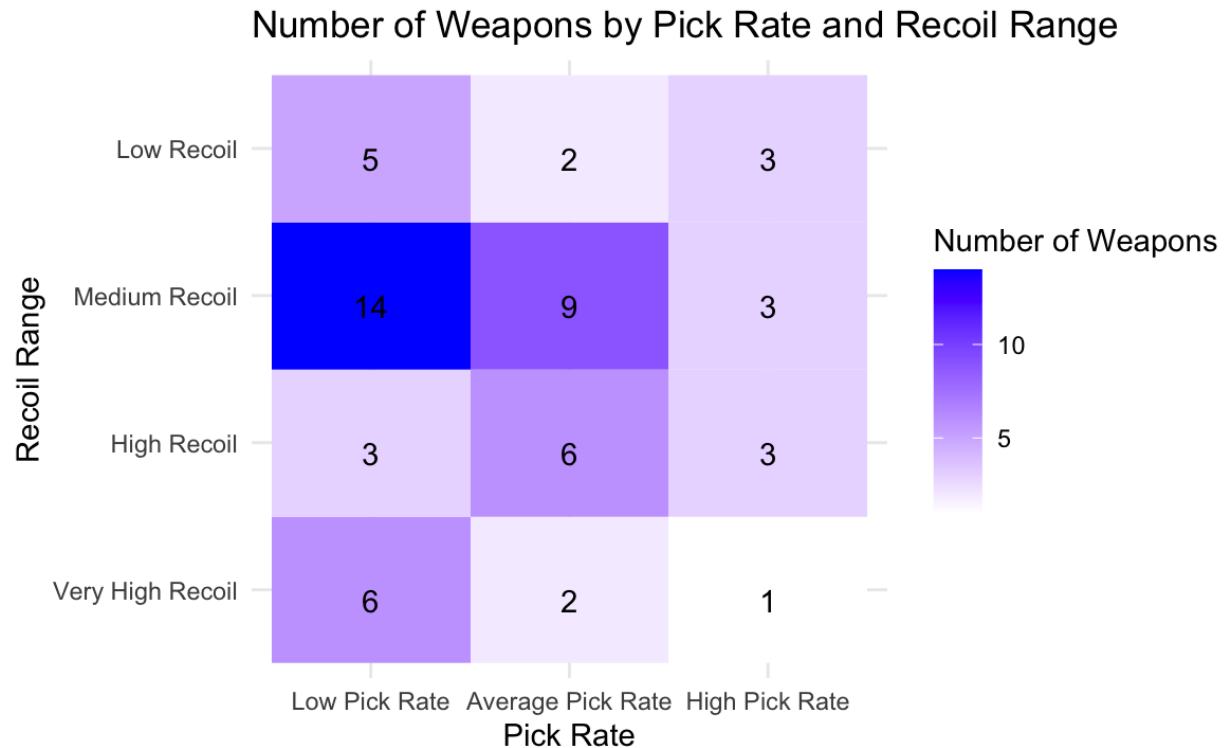
ggplot(low_pick_df, aes(x = Weapon, y = Recoil)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  coord_flip() +
  labs(title = "Recoil of Low Pick Rate Weapons",
       x = "Weapon",
       y = "Recoil Value") +
  theme_minimal()
```



This is a measure of spread that informatively shows a somewhat - mixed distribution with low-pick weapons recoil scores being mostly skewed around 0 but with more negative values than the high pick rate weapons.

```
# Count the number of weapons in each combination of Pick Rate and Recoil
range
heatmap_data <- as.data.frame(table(fulldf$PickRate, fulldf$RecoilRange))
colnames(heatmap_data) <- c("PickRate", "RecoilRange", "Count")
# Reorder the Levels of PickRate
heatmap_data$PickRate <- factor(heatmap_data$PickRate, levels = c("Low Pick
Rate", "Average Pick Rate", "High Pick Rate"))

# Plot
ggplot(heatmap_data, aes(x = PickRate, y = RecoilRange, fill = Count)) +
  geom_tile() +
  geom_text(aes(label = Count), vjust = 1) +
  scale_fill_gradient(low = "white", high = "blue") +
  labs(title = "Number of Weapons by Pick Rate and Recoil Range",
       x = "Pick Rate",
       y = "Recoil Range",
       fill = "Number of Weapons") +
  theme_minimal()
```



Statistical Test:

```
# Ensure PickRate is an ordered factor
fulldf$PickRate <- factor(fulldf$PickRate, levels = c("Low Pick Rate",
"Average Pick Rate", "High Pick Rate"), ordered = TRUE)

# Convert to numeric for correlation test
fulldf$PickRateNumeric <- as.numeric(fulldf$PickRate)

# Use Kendall's Tau for correlation
cor.test(~ Recoil + PickRateNumeric, data = fulldf, method = "kendall")

##
## Kendall's rank correlation tau
##
## data: Recoil and PickRateNumeric
## z = 0.007991, p-value = 0.9936
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0.0009311689
```

In analyzing the relationship between Recoil (a continuous variable) and PickRate (an ordinal variable), I employed Kendall's Tau correlation test. The choice of Kendall's Tau was due to its ability to assessing associations involving at least one ordinal variable.

The resulting p-value from this test was greater than the conventional alpha level of 0.05. This basically means that the evidence is not sufficient enough to reject the null hypothesis of *no association* between Recoil and PickRate. I cannot assert that changes in Recoil scores are consistently related to changes in PickRate categories.

However, it is important to note that failing to reject the null hypothesis **does not equate to accepting it**. The absence of significant evidence is not the same as evidence of absence. Therefore, we cannot **conclusively state that there is no relationship between Recoil and PickRate; only that our study did not find sufficient evidence of such a relationship under the conditions tested**.

Regarding the use of Kendall's Tau, given that Recoil is a continuous variable and PickRateNumeric is an ordinal encoding of PickRate, the test is appropriate and meets assumptions. Kendall's Tau is designed to handle such types of data, making it a suitable choice for this analysis.

A power analysis would help determine if the study had a sufficient sample size to detect an effect, if one exists. A low power could mean that the study might not be able to detect a small but potentially meaningful association. Therefore, interpreting these results should be done in the context of the study's power, alongside the considerations of the test's assumptions and the data's characteristics.

## Limitations

In reflecting on the results of the Kendall's Tau correlation test and the p-value exceeding 0.05, it's crucial to consider the limitations of the dataset and how they might influence the findings. The categorical nature of the Pick Rate column, with only three levels, limits the granularity of the analysis. A more detailed quantitative measure of Pick Rate, such as exact percentages of operator picks per number of rounds in a given number of matches, could potentially reveal trends that the current categorical data obscures.

Similarly, the Recoil scores in my dataset lack the depth and detail available to Ubisoft. The omission of additional attachments that affect recoil is a significant limitation, as these could have a considerable impact on the weapon's performance and, consequently, its pick rate. This limitation highlights the potential for more nuanced and comprehensive data to yield different insights.

For future research directions, considering 'Ban Rates' of operators in relation to their weapon's recoil could offer another valuable perspective. This approach might uncover correlations between high ban rates and specific weapon characteristics, offering a different angle from which to understand player preferences and game balance.

Such investigations would benefit from more advanced statistical techniques like factor analysis and mixed-effects models. However, the current dataset, with its non-continuous and non-numeric features, is unsuitable for these methods. Additionally, the potential for multicollinearity, given that weapon attributes are often designed to balance each other,

poses a challenge. High damage, for example, might be counterbalanced by high recoil, making it difficult to disentangle these factors.

In light of these limitations, a power analysis is particularly relevant. It would help in assessing whether the current study had enough power to detect a small but meaningful relationship, if one exists. The absence of significant findings in this study might be due to insufficient data detail or sample size, rather than a true lack of correlation. This consideration underscores the need for a more comprehensive dataset and a robust analytical approach to more accurately evaluate the relationship between weapon characteristics and player choices in the game.