

## **Examining Gender and Provocation Effects on Aggression in Experimental Data**

HUDM 5001 Programming for Data Science

Teachers College, Columbia University

## Abstract

This study examined gender differences in trait aggression and behavioral aggression under varying levels of provocation. Aggression was measured using the Buss–Perry Aggression Questionnaire (BPAQ) and a provocation-based behavioral task with three levels of increasing provocation. Independent-samples t-tests and two-way ANOVA were used to examine gender and situational effects. Results showed no significant gender difference in overall BPAQ scores. However, behavioral aggression increased across provocation levels for both genders, with no significant interaction between gender and provocation. These findings suggest that situational provocation plays a stronger role in aggressive behavior than gender-based differences in trait aggression, while also highlighting the important contribution of individual differences in trait aggression to aggressive responding.

## Introduction

Aggression is a topic that has been studied across many disciplines because it can appear in everyday interactions as well as in controlled laboratory tasks. Researchers generally agree that aggression does not come from a single cause. Instead, it often reflects a combination of (1) something in the situation that triggers a reaction, (2) personal tendencies that make someone more sensitive to that trigger, and (3) how well a person can manage or hold back their impulses. A simple way to describe this idea is the Instigation–Impellance–Inhibition Model, which basically suggests that people react based on what sets them off, how strongly they are pushed toward reacting, and how effectively they can regulate themselves.

Instigation refers to the kind of event that sets off a reaction—for example, criticism, frustration, or feeling provoked. Impellance captures personal traits that make someone more likely to react strongly. Several studies show that characteristics like perfectionism or shyness can increase emotional sensitivity, making people more reactive when facing stress or self-evaluation (Martinez et al., 2022; Sun et al., 2021). Inhibition describes a person’s ability to regulate or hold back their reactions. For example, some people respond immediately when frustrated, while others pause and restrain themselves. Some research also notes that gender norms can influence how openly aggression is expressed, with males sometimes showing more outward behavioral responses (Gómez-Leal et al., 2024).

This general idea fits well with the dataset we are analyzing. The questionnaire scores represent relatively stable tendencies (similar to “impellance”), while the block-level provocation task introduces the “instigation” that pressures participants in real time. How aggression changes across the blocks may also relate to individual differences in inhibition, especially since emotional regulation is often linked to aggressive behavior in experimental tasks (Hsieh & Chen, 2017).

The dataset follows a structure that is consistent with the studies described in the related literature, where both self-report measures and behavioral tasks are used together to understand how general tendencies relate to reactions under pressure. Some research has shown that emotional regulation difficulties or stress-related traits can increase aggressive responding in structured tasks (Zeyrek-Rios et al., 2025). Because gender differences are frequently mentioned

across both questionnaire-based and laboratory aggression studies, it is reasonable to examine whether similar patterns appear in this dataset as well.

Based on the theoretical framework described above, the present study examines the relationship between gender, trait aggression, and situational provocation. Specifically, the project addresses two research questions. Research Question 1 (RQ1) asks: How is gender related to self-reported aggression as measured by the Buss–Perry Aggression Questionnaire (BPAQ)? This question addresses whether males and females differ in their overall levels of trait aggression. Research Question 2 (RQ2) asks: How are gender and provocation level related to behavioral aggression scores? This question focuses on how aggressive responses change from different provocation levels, and whether males and females show different patterns of behavioral aggression as provocation intensifies.

### Data

This dataset originates from the study by Hsieh and Chen (2017) published in PLOS ONE. The dataset used for this project includes three main components: participants' gender, self-reported aggression scores from the Buss–Perry Aggression Questionnaire (BPAQ), and behavioral aggression scores collected during a provocation-based task. This layout is similar to the experimental paradigms described in aggression research, where participants first report their general attitudes or tendencies and then respond to a situation that becomes increasingly provocative (Hsieh & Chen, 2017).

**Table 1**

*Descriptive Statistics for Variables Related to the Research Questions*

Variable	Type	Operational definition
Gender	Categorical	Male / Female
BPAQ total score	Continuous	Self-reported trait aggression (BPAQ total)
Block 1 mean	Continuous	Mean aggression under low provocation
Block 2 mean	Continuous	Mean aggression under medium provocation
Block 3 mean	Continuous	Mean aggression under high provocation

These three block means show how participants respond as provocation increases. This structure makes it possible to examine whether aggression rises in response to the escalating situation, and whether males and females follow similar or different patterns. Prior findings suggest that individuals who have difficulty with emotional regulation or who experience stronger negative reactions tend to show higher aggression as situational stress increases (Zeyrek-Rios et al., 2025), which makes this dataset suitable for exploring such patterns.

This dataset works well for the goals of the project. First, it includes both questionnaire and behavioral measures, allowing us to compare internal tendencies with actual reactions. Second, the provocation task is clearly divided into three levels, making it possible to observe changes over time. Third, Gender is a meaningful grouping variable supported by many prior

findings on aggression (Gómez-Leal et al., 2024). Finally, the dataset is clean and manageable, fitting the scope of the course and the timeline of the project

Data preparation was conducted on the dataset of 80 observations to ensure its statistical suitability. A systematic check for missing values was conducted across the five relevant variables. The output confirmed zero missing entries for all columns. As a result, no imputation or listwise deletion of cases was required, ensuring the initial sample size of 80 observations was maintained for the subsequent cleaning steps. Thereafter, the data type of the gender column was explicitly converted to category variable using the `.astype()` function. This adjustment ensures the variable is correctly classified as nominal (a grouping variable) for statistical analysis. The data cleaning steps involved outlier detection and removal. Outliers were systematically identified using the Interquartile Range (IQR) method. This method calculated the  $1.5 \times \text{IQR}$  bounds for each of the four continuous score columns. Any observations falling outside this range were iteratively removed from the dataset. This procedure resulted in the exclusion of 8 rows, yielding a final dataset (`df_clean`) of 72 observations.

Data preparation steps were taken to establish a reliable foundation for analysis. Checking for missing data confirms sample completeness, while adjusting the gender type ensures correct statistical model specification. Outlier removal is crucial because extreme scores significantly distort central tendency and variance estimates, which can violate the assumptions of parametric tests (t-tests and ANOVA in our study), thereby compromising the validity and reliability of the findings.

### **Data Processing Methodology**

Research Question 1 examines the relationship between gender and self-reported aggression as measured by the Buss–Perry Aggression Questionnaire (BPAQ). Boxplots were first used to visually compare the distributions of BPAQ total scores between males and females. Subsequently, an independent-samples t test was conducted to formally test whether there was a statistically significant difference in BPAQ scores between genders.

Research Question 2 investigates how gender and provocation level are related to behavioral aggression scores, line plots were used to visualize changes in mean aggression across three levels of provocation (low, medium, and high) separately for males and females. To statistically evaluate these patterns, a two-way analysis of variance (ANOVA) was performed, with gender and provocation level as factors. This analysis assessed both the main effects of provocation and gender, as well as their interaction.

Finally, to further explore individual-level relationships between trait aggression and behavioral aggression, a faceted scatterplot with regression lines was used. This visualization illustrated how BPAQ total scores were associated with aggressive responses across different provocation levels and genders, providing a descriptive examination of potential interaction patterns at the individual level.

## Results

### Descriptive analysis

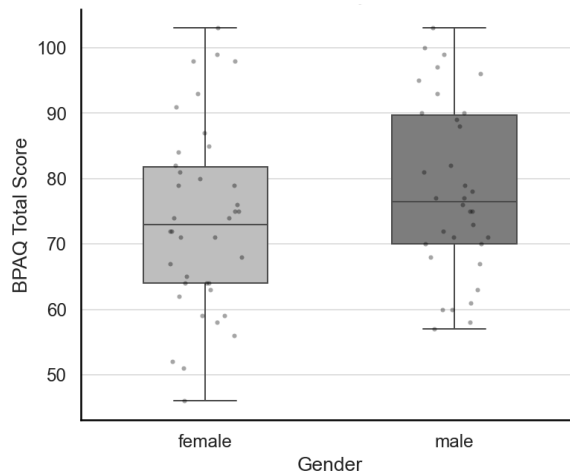
The descriptive statistical analysis was conducted on the cleaned dataset (df\_clean) to examine the distribution of the sample and the core aggression scores across gender and blocks. The final sample used for analysis was balanced, with equal numbers of participants in both groups - Female (N = 4) and Male (N = 4). The descriptive statistics for the Buss-Perry Aggression Questionnaire (BPAQ) total scores demonstrated a clear difference in self-reported aggression between the two groups - Female (Mean = 50.75, SD = 4.35) and Male (Mean = 64.25, SD = 4.35). The overall mean aggression scores across all participants showed a consistent upward trend across the three Blocks - Block 1 (Mean = 5.00), Block 2 (Mean=5.84) and Block 3 (Mean=6.74). The results show two key patterns. First, the Male group reported a substantially higher mean total score on the BPAQ (64.25) compared to the Female group (50.75). This suggests that the male participants, on average, self-reported higher baseline aggression. Second, a distinct escalation effect was observed across the experimental conditions: the mean aggression score consistently increased from Block 1 (5.00) to Block 2 (5.84) and further to Block 3 (6.74). This pattern suggests that the experimental manipulation across the blocks led to increasingly aggressive responses over time.

### Data visualization and statistical testing

Research Question 1 tried to investigate whether there was any difference between males and females in their total scores on the Buss-Perry Aggression Questionnaire (BPAQ).

**Figure 1**

*Gender Differences in Self-Reported Aggression (BPAQ Total Scores)*



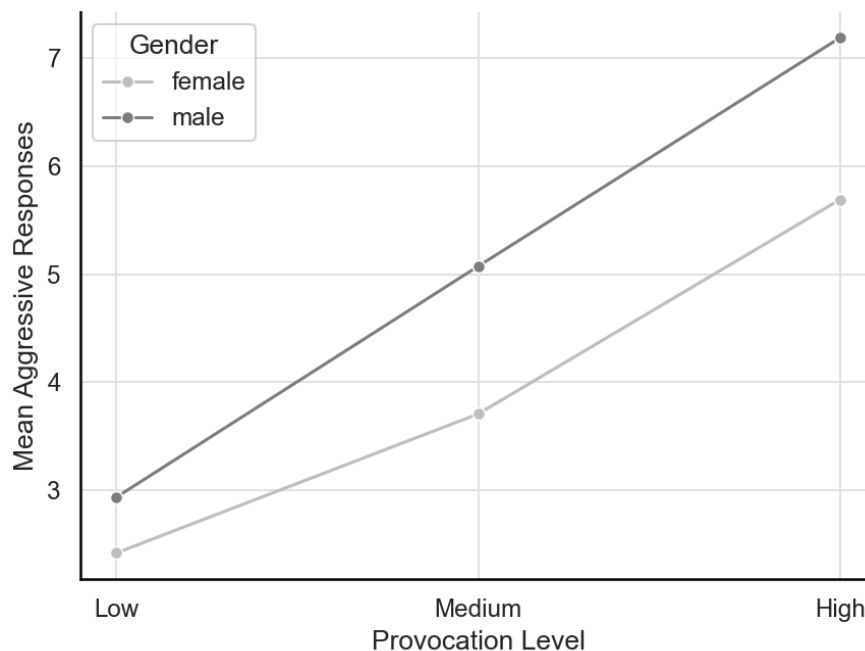
We first use a box plot to visualize the distribution of BPAQ scores for different genders. As shown in Figure 1, the median BPAQ total scores for male and female participants were broadly similar, although the median appeared slightly higher for males than for females. The distributions showed substantial overlap, the distribution of BPAQ scores among females appeared more dispersed.

We further conducted hypothesis testing using an independent sample t-test to assess whether there are significant differences in the mean BPAQ scores between the two gender groups. The null hypothesis ( $H_0$ ) stated that there was no difference in BPAQ total scores between males and females, whereas the alternative hypothesis ( $H_a$ ) posited that a difference did exist. After establishing these hypotheses, participants were grouped by gender into male and female categories to facilitate subsequent statistical analyses. The analysis indicated a small mean difference between genders, with males scoring on average 4.66 points higher than females on the BPAQ. Even with a sufficient difference between groups, the results still did not reach statistical significance ( $t(70) = 1.434, p = .156$ ). Since the p-value exceeded the significance level of  $\alpha = .05$ , we failed to reject the null hypothesis. In summary, there was insufficient statistical evidence to conclude that BPAQ scores differed significantly by gender.

Research Question 2 was a further extension of Research Question 1 and tried to examine whether there was any difference in aggression scores across blocks and whether this pattern differed between males and females.

**Figure 2**

*Gender Differences in Behavioral Aggression Across Provocation Levels*



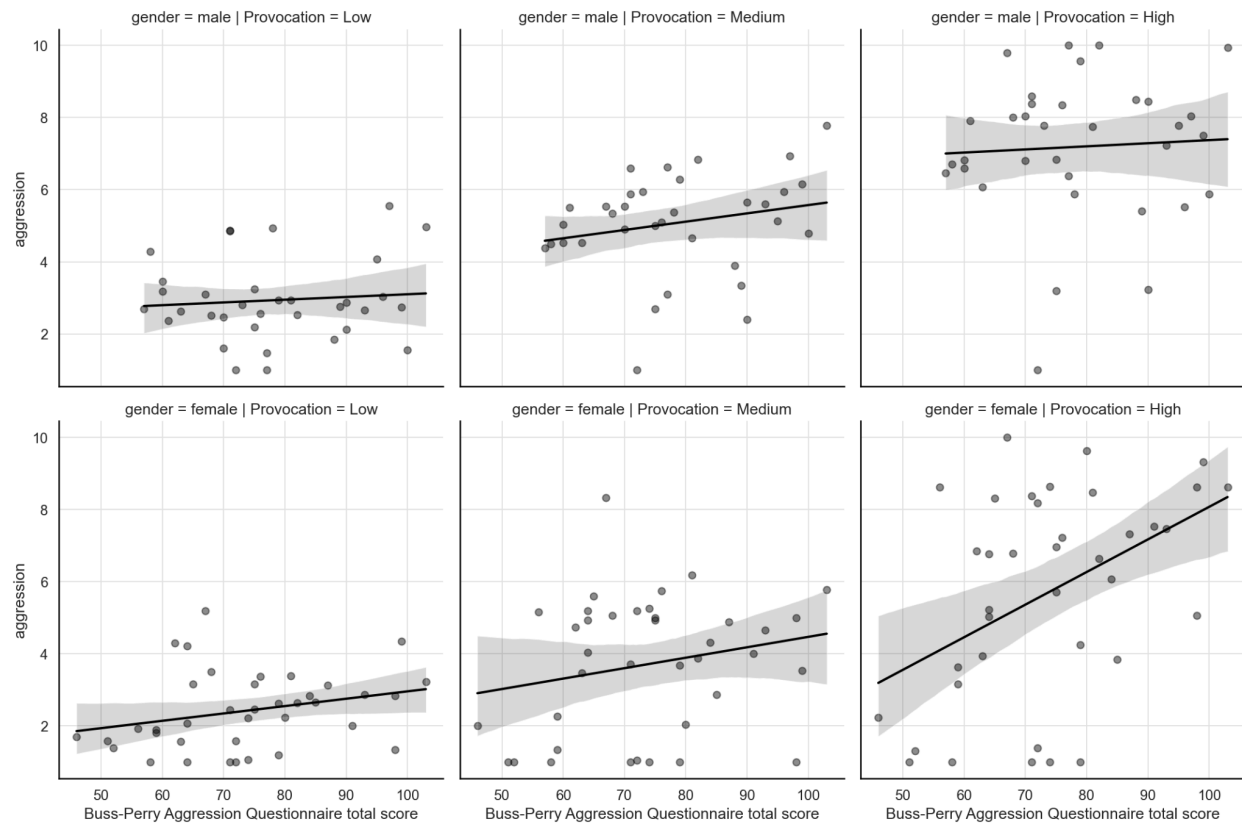
To address this question, we first used a line plot to visualize the scores of different genders under different provocation conditions. Figure 2 shows that aggressive behavior in both genders increases with the level of provocation. Furthermore, men's scores were higher than women's across all provocation scenarios.

Then, we use a two-way analysis of variance (ANOVA) was conducted to examine the effects of gender, block, and their interaction on aggression scores. The results show that aggression scores were different across blocks, with Block 1 having the lowest average score and Block 3 having the highest average score. The ANOVA results indicated there was a significant

main effect of block on aggression scores(  $F(2, 210) = 73.38, p < .001$  ),given that aggression levels means were significantly different across the three blocks. There was also a significant main effect of gender on aggression scores(  $F(1, 210) = 19.96, p < .001$  ), indicating an overall difference in aggression levels between males and females. However, the interaction between gender and block was not significant(  $F(2, 210) = 1.49, p = .228$  ), indicating similar patterns of aggression across blocks for males and females.

**Figure 3**

*Individual-Level Relationship Between BPAQ and Aggression Across Provocation Levels and Gender*



Faceted scatterplot with regression lines (Figure 3) further illustrates the gender differences in aggressive behavior exhibited by individuals with different levels of trait aggression under different provocation scenarios. Consistent with previous research findings, both males and females, and individuals with varying levels of trait aggression, showed higher levels of aggressive behavior as the level of provocation increased. However, there appeared to be gender differences in how trait aggression manifested in aggressive behavior in specific situations. The relationship between BPAQ scores and aggressive behavior across genders became more significant as the level of provocation increased. Under high provocation conditions, women with high trait aggression showed a greater increase in aggressive responses than men, while at low levels of provocation, the difference between trait aggression and aggressive behavior was less pronounced. These patterns are merely descriptive results based on

visualized data; the exact results should be determined through statistical tests for interaction effects.

Overall, the results from Research Question 1 and Research Question 2 both help us clarify the relationship between aggression, gender, and situational factors in this study. The results of Research Question 1 showed that there was no significant difference between males and females in Buss–Perry Aggression Questionnaire (BPAQ) scores. In contrast, the results of Research Question 2 indicated significant differences in aggression scores across blocks, with aggression increasing as provocation level increased. Figure 3 provides individual-level evidence suggesting that the behavioral expression of trait aggression may vary by gender under specific situational conditions. While males and females responded similarly to increasing provocation on average, the visualized data indicate that under high provocation, females with higher BPAQ scores tended to show a steeper increase in aggressive responses than males. These results suggest that situational factors play an important role in changes in aggressive behavior. Although males and females differed in their overall average aggression levels across the block-based task, the interaction between gender and block was not significant. This indicates that, despite gender differences, the pattern of aggression change across blocks was similar for both groups.

The findings suggest that when examining aggressive behavior, it is important to consider both individual characteristics and situational stressors, rather than focusing only on gender.

### **Conclusion**

This project examined how gender and situational provocation were related to both self-reported and behavioral aggression using data from the BPAQ questionnaire and a three-block provocation task. Across our analyses, several patterns emerged that clarify the roles of individual traits and situational factors in shaping aggressive responses.

For self-reported aggression (RQ1), the independent-samples t-test showed no significant gender difference in BPAQ total scores. Although males had slightly higher mean scores than females, the gap was not large enough to be meaningful. This suggests that, in this sample, males and females reported similar levels of trait aggression. In contrast, the provocation-based behavioral task (RQ2) revealed strong situational effects. Aggression increased steadily from Block 1 to Block 3, indicating that participants became more aggressive as provocation intensified. ANOVA confirmed a significant main effect of provocation level. Gender also showed a main effect, with males displaying higher aggression overall, but the gender by provocation interaction was not significant, meaning both groups showed similar escalation patterns.

The individual-level visualization provided additional nuance by suggesting that trait aggression may relate differently to behavior depending on gender and provocation intensity. Under high provocation, females with higher BPAQ scores appeared to show a steeper rise in behavioral aggression than males. While this pattern was not tested statistically, it points to potential directions for future work.



In conclusion, the findings highlight the importance of considering both personal characteristics and situational triggers when studying aggression. Beyond this specific dataset, the results may be useful for researchers or educators interested in understanding how individuals respond under increasing pressure, or for those designing studies that combine trait measures with real-time behavioral tasks. From a programming perspective, future improvements could include automating additional statistical comparisons, expanding the Python classes to handle repeated-measures designs, or integrating more flexible visualization functions. With more time, the program could also be extended to allow interactive exploration of individual trajectories, making the analysis more adaptable for larger or more complex datasets.

## References

- Gómez-Leal, R., Fernández-Berrocal, P., Faria, L., Costa, A., & Cabello, R. (2024). *The role of emotional intelligence and gender in the relationship between implicit theories of emotions and aggression: Moderated mediation model in young individuals*. International Journal of Adolescence and Youth, 29(1), 2363352. <https://doi.org/10.1080/02673843.2024.2363352>
- Hsieh, I.-J., & Chen, Y. Y. (2017). *Determinants of aggressive behavior: Interactive effects of emotional regulation and inhibitory control*. PLOS ONE, 12(4), e0175651. <https://doi.org/10.1371/journal.pone.0175651>
- Martinez, K., Guan, C., & Walker, D. C. (2022). *I'm angry that I'm not perfect! Aggression mediates the relationship between perfectionism and eating pathology in men and women*. Eating and Weight Disorders – Studies on Anorexia, Bulimia and Obesity, 27, 1895–1901. <https://doi.org/10.1007/s40519-021-01319-w>
- Sun, L., Fu, Z., & Zheng, Y. (2021). *Shyness and loneliness in Chinese young adults: Roles of aggression and gender*. Journal of Aggression, Maltreatment & Trauma, 30(1), 43–53. <https://doi.org/10.1080/10926771.2020.1725209>
- Zeyrek-Rios, E. Y., Çetinöz, E., Erten, I., Gürsoy, A., & Gönültaş, M. B. (2025). *Relationship between emotion regulation and aggression mediated by criminal thinking in women convicted of violent offenses*. Women's Studies International Forum, 113, 103179. <https://doi.org/10.1016/j.wsif.2025.103179>