Chess's Impact on Memory

Alec Brooks

2025-02-24

Abstract

This study explores the relationship between playing chess and memory improvement using data from a simulated population on The Islands platform. Over a simulated fall semester, students were divided into control and participant groups, with the participants playing daily chess games while both groups completed daily memory tests. Although the participant group showed greater improvement in memory scores compared to the control group, the difference was low statistical significance, likely due to the small sample size. However, linear regression revealed that study duration and chess win records were meaningful predictors of memory scores, accounting for a substantial portion of the observed variation. These findings suggest a potential link between playing chess and memory enhancement, but further research with larger, real-world populations is necessary to confirm these results.

Introduction

Chess is a game so old its origins have passed into legend, shrouded in mystery that has captured imaginations for centuries. This historical allure often lends the game an air of exclusivity, causing people to associate it with the smartest individuals and the assumption that excelling at chess requires innate intelligence. However, many modern chess masters argue that mastery of the game is not rooted in intelligence but rather in memory. As the late chess Grandmaster Bobby Fischer famously stated, "A strong memory, concentration, imagination, and a strong will is required to become a great chess player."

This observation raises an intriguing question: Is good memory a prerequisite for mastering chess, or does playing chess enhance memory capacity over time? This study seeks to explore this relationship by observing students over the course of their fall semester, using simulated data provided by The Islands website. Specifically, we aim to determine whether regular engagement with chess can improve overall memory in young adults.

Methodology

The methodology of this study is divided into two phases: Data Gathering and Analysis.

Data Gathering

The data collection phase began with the random selection of 12 students distributed across the three islands of The Islands simulation. A random sampling method was employed to select four students from each island, ensuring equal representation. Within each island, two students were assigned to the control group, while the remaining two formed the participant cohort. While a larger sample size would have been ideal for a more robust analysis, logistical and time constraints necessitated a smaller sample.

The testing phase spanned the entire simulated fall semester. During this period, all students took a daily memory test using a consistent 10-card memorization format. The participant cohort also played two games of chess each day—one as Black and one as White—with their wins and losses recorded.

Data Analysis

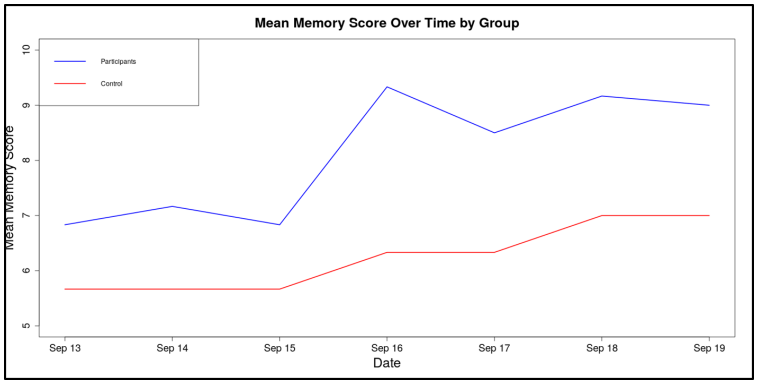
The analysis was conducted using R, focusing on the relationship between regular chess playing and daily memory test scores. The Date variable, representing the duration of the study, was a primary focus. Initially, an F-test was conducted to ensure equality of variances before proceeding with a Welch Two Sample t-test. Following the t-test, linear regression was used to examine the impact of time and chess game outcomes on memory scores. Lastly, predictions were made to estimate memory scores if the study were extended beyond its original duration.

Data Description'

The data used in this study was recorded from simulated tests performed on the inhabitants of The Islands, a platform created by the University of Sydney, Australia. The Islands hosts a simulated population of families living on three culturally distinct islands, enabling students to design experiments and explore research methodologies. Key variables include the Date of data collection, Student Name for identification, and School. There is a Participant variable that indicates whether a student actively participated in the chess-playing intervention (True) or was part of the control group (False), while the Memory Test variable measures cognitive performance. Two additional variables, as_Black and as_White, categorically represent outcomes of chess games played, with -1 indicating a loss, 1 a win, and 0 no participation. Although the data reflects experiments conducted, it is based on a simulated population. Thus, the findings should not be generalized to real-world settings but rather serve as a demonstration of methodologies that can be applied to similar research questions. All results are specific to the population of The Islands. The dataset is publicly available at the following link:

https://www.kaggle.com/datasets/alecbrooks/the-islands-chess-experiment

General Improvement & Statistical Significance



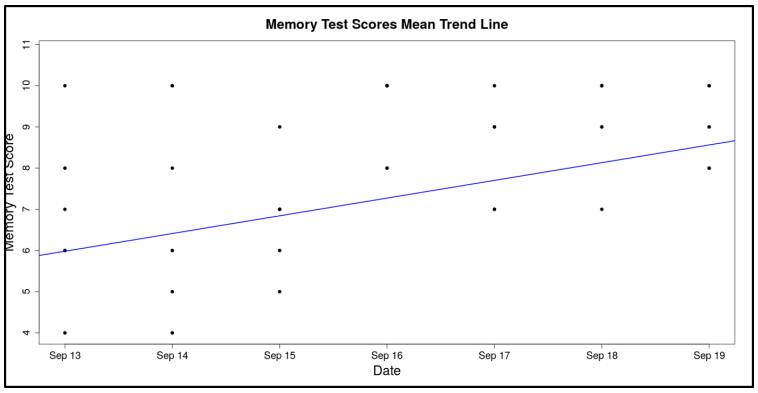
Examining a time series of both the participant and control groups over the course of the study shows notable differences in the starting and ending average memory test scores. While both groups demonstrated improved memory scores, the participant group's overall improvement significantly outpaced that of the control group. The table below highlights the mean improvement from the start to the end of the study, showing that the participant cohort's memory scores increased by almost twice as much as those of the control group.

Table 1: Table 1: Mean Memory Score Improvement

Group	Mean
Participants	2.17
Control	1.33

A Welch Two Sample t-test on mean memory score improvement yielded a p-value of p=0.197, which is not statistically significant. This may be due to the small sample size of only 12 students, leading to reduced accuracy. While the relatively low p-value suggests potential, a larger sample size would be necessary for more reliable findings. Given this, we will proceed with further exploration through linear regression to identify trends, but all results must be interpreted cautiously due to the statistical uncertainty.

Linear Regression and Further Exposure



An initial linear regression using the date of the study as the sole predictor variable for memory test scores yielded a statistically significant p-value (p=0.00100). This result indicates that the length of participation in the study is a significant predictor of memory test scores. However, the R squared value was low (0.2396), suggesting that the date of the study alone accounts for only about 24% of the variation in memory test scores.

By extending the model to a multiple regression that includes the win-loss record for games played as Black and White, we observe a meaningful improvement. Adding these variables increased the R squared to 0.3487, indicating that the model now accounts for approximately 35% of the variation in memory test scores. This model also yielded a statistically significant p-value (p=0.0008963), demonstrating that longer involvement in the study, coupled with more games won, had a significant impact on memory test scores within the participant cohort.

With this updated model, we can make predictions about how memory test scores might change if the study were extended by three additional days. By focusing on the Date variable, we can estimate the average memory score for participants during this extended period.

In the predictions, the fit value represents the estimated memory test score, while lwr and upr indicate the lower and upper bounds of the confidence interval, showing the possible variation in the prediction.

Table 2: Table 2: Prolonged Study Prediction

Date	fit	lwr	upr
2024-09-20	7.97	7.01	8.93
2024-09-21	7.97	7.03	8.91
2024-09-22	8.41	7.60	9.23