An Analysis of the Success of Chess Opening Principles in Different Rating Brackets

Thomas Benoit, Joshua Goubeaux

I. Introduction

In the game of chess, strong players typically follow established opening principles. These principles are broad-level strategies that lead to strong play early-on within the game. For instance, the strategy of controlling the center with your pieces rapidly is an example of a common principle many player employ in their games. Players who follow this strategy in the opening typically find themselves with enough board control to safely develop their other pieces from their original squares. Players who fail to follow this principle tend to struggle developing their pieces onto useful squares, resulting in a worse position.

For the purpose of this study, we will be quantifying some of the most common principles and analyzing millions of games to determine which ones have the highest impact on winning chances. Coaches and players base what strategies "work" against their opponents based off experience and recommendations from other players. But which strategies are *actually* the most successful?

Chess coaches typically recommend different principles to players with different ratings. What may be ideal for one player at a higher rating may be less favorable for a player in a lower rating. For example, many coaches recommend fundamental chess principles to beginners because fundamental strategies tend towards positions that are easier for beginners to play. Other principles tend to perform better in different rating brackets because of their tricky nature. For instance, so-called "gambit" openings (wherein one side sacrifices material early-on to gain a lead in piece development) tend to be most successful in the moderate level games. This is because many players are skilled enough to handle such an opening yet their opponents may not be experienced enough to refute the risky strategy. It is for these reasons that we are investigating the success of principles within the scope of distinct rating brackets.

This paper details the processing and evaluating of opening principles in over one million chess games. We measured the success of four different strategies across four distinct rating brackets :

1. Castle early to get your king to safety.

2. Gain early control of the center four squares (d5,e5,d4,e4) early on.

3. Play a gambit opening (give up material early on for accelerated piece development)

4. Develop multiple pieces instead of moving the same piece multiple times in the opening.

We hypothesised that all of these strategies will correlate with higher winning chances. We also predict that these opening principles will be most successful at the mid-level rating brackets. This is because low-level games are typically full of blunders that heavily impact the game and high-level games are usually determined by more minute positional details.

II. Related Work

Historically, modern computational chess analysis has a heavy bias towards either finding the best moves or evaluating positions. This type of analysis is very useful for exploring specific situations especially with the speed of modern computers. Through the use

of modern engines, players can perform near-perfect analysis of any game within a matter of seconds. Because of the dominance of chess engine development, little research has been conducted on analyzing databases of games.

Some limited research on automatic commentary constructions has been conducted in the past 40 years [1]. These constructions annotate games by generating useful commentary that goes beyond the strength of a move while typical chess engines only comment on the strength of a move. One of the earliest breakthroughs in this topic was in 1993 when the ICCA Journal hosted a competition for the best automated chess commentary program. A program dubbed "Chessmaster 4000" won this competition in 1993 [2]. The Chessmaster series went on to sell millions of copies of their automatic commentary construction [3]. Although commentary programs such as the Chessmaster 4000 are similar to engines as they perform analysis of in-progress games, they share similarities with our post-game opening principle analysis. The commentary systems perform feature analysis to extract high-level strategies comprehensible to humans. We performed similar feature analysis to a large database of games when we extracted and analyzed the success of opening principles and strategies.

This approach of analysing features midgame is useful for allowing players to receive meaningful analysis in real-time. The main drawback of this method is the lack of analysis involving multiple games. Our program performs analysis of chess principles across many games.

III. Solution

With the ultimate goal of determining which chess principles contribute most to the outcome of a game, we decided that the best solution was to treat each principle as a feature acquired from individual games. Each feature could then be labeled with the game's outcome, forming a data set.

Using this newly created data set, we then found the most influential features using univariate feature selection. In performing this analysis, we could confidently verify the influence any specific feature has on the outcome of a game. Since these features each correlate to a chess principle, we found the influence of each principle from our approach.

IV. Data

One of the great advantages of online chess is the abundance of accurately recorded games. We used games from the popular online chess site lichess.org. Lichess is an open-source chess platform that records all rated games played on their website and releases them for free on database.lichess.org. Since these games were played online, there are no illegal moves or errors in the recorded games. Currently, their database contains over 2 billion online chess games played on their site. The games are stored in a format called portable game notation (PGN), meaning each game is represented by a short section stating general information about the game followed by a list of in-order moves that were played. Figure 1 shows an example of a popular game saved in PGN format. Each PGN, when uncompressed, is about 0.8 kilobytes in size on average. Because of hardware limitations, we selected a sample of 1.6 million games to perform our analysis on. This sample selection of games was 1.2 GB in size.

Recall chess is a two player game where each game is a collection of moves played by two players. Although we only sampled 1.6 million games, we extracted features for both white and black, resulting in a data base of 3.2 million entries where each entry represents a player's adherence to opening principles and their success or failure in the game.

V. Setup

While the quantity of data available was quite substantial, it became difficult to parse the features in its raw form. As a result, before any features could be extracted, the data was transformed from its PGN format

```
[Event "Paris"]
[Site "Paris FRA"]
[Date "1858.??.??"]
[EventDate "?"]
[Round "?"]
[Result "1-0"]
[White "Paul Morphy"]
[Black "Duke Karl / Count Isouard"]
[ECO "C41"]
[WhiteElo "?"]
[BlackElo "?"]
[PlyCount "33"]

1.e4 e5 2.Nf3 d6 3.d4 Bg4 4.dxe5 Bxf3
5.Qxf3 dxe5 6.Bc4 Nf6 7.Qb3 Qe7 8.Nc3 c6
9.Bg5 b5 10.Nxb5 cxb5 11.Bxb5+ Nbd7
12.O-O-O Rd8 13.Rxd7 Rxd7 14.Rd1 Qe6
15.Bxd7+ Nxd7 16.Qb8+ Nxb8 17.Rd8# 1-0
```

Fig. 1: Example of the PGN format (of the famous Opera Game)

into a much more readable comma-separated values (CSV) format, containing a list of moves performed by each side and the game's winner. During this process, games were separated into four different CSV files based on the player's rankings: 0-999, 1000-1499, 1500-1999, 2000+. By separating the games into these different brackets, the games could later be analyzed to determine whether a certain play style was more beneficial to people of different skill levels.

Once the data was converted to a more suitable form, the features could be extracted from each side. We based our features on many of the popular opening principles most chess players value in their own games. In order to quantify these principles, we had to come up with a measure of adhesion to each principle.

Table I: Principles and their measures

_	Principle	Measure
1	Castle early	The turn a player castles
2	Gain early control of the center	The number of center squares controlled by move five
3	Play a gambit opening	The material difference in pieces by move four.
4	Develop multiple pieces early on	The number of unique pieces moved by move seven.

Each set of features was then labeled with if the player won the game or not, thus completing the data set. The final feature data set had entries of the form "when_castled, num_center_squares_controlled, material_ difference, num_unique_pieces_moved, win or loss"

```
game = game in PGN format
For all games in database:
    move_list, winner = preprocess(game)
    add_to_rating_range_file(move_list,
        winner)

for each rating_range_file:
    for each game in file:
        features[] = extract_features()
        add_whites_game(features,win/loss)
        add_blacks_game(features,win/loss)

For each rating_range_features_file:
    analyze_all_games(features,win/loss
        ,...,features,win/loss)
```

Fig. 2: Pseudo code of the data extraction and analysis

To identify the importance of each feature, we used univariate feature selection, acquired from sklearn's SelectKBest function. Univariate feature selection is typically used to determine the most poignant features in a data set and prune the less impactful features. This is performed to reduce noise when training a model. Since univariate selection assigns a numerical importance score for each feature in the data set, we can us these numbers to determine which features are relevant towards our target and which aren't. Because SelectKBest scores features on a particular scoring functions, we had to pick the best scoring function to use with our data. We chose to use the f_classif function which calculates the ANOVA F-value for each feature of the provided sample. From these F-values, SelectKBest returns the four highest values. Since our data set only has four features, we were able to use these values to determine the success of each opening principles. We are assuming that features that have a high impact on the outcome of the game means they are successful.

VI. Results and Analysis

Patterns begins to emerge when examining each feature's importance at each rating range. The castling move number holds the highest importance among all ratings, followed by the number of center squares controlled, and then gambits. Castling also appears to have a right-scewed distribution across the rating ranges.

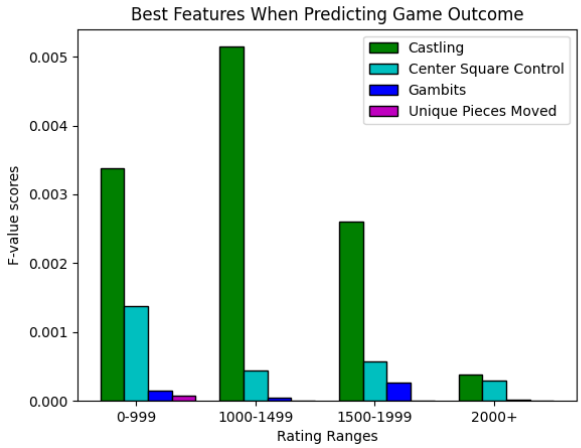


Fig. 3: Feature Importance When Predicting Game Outcome

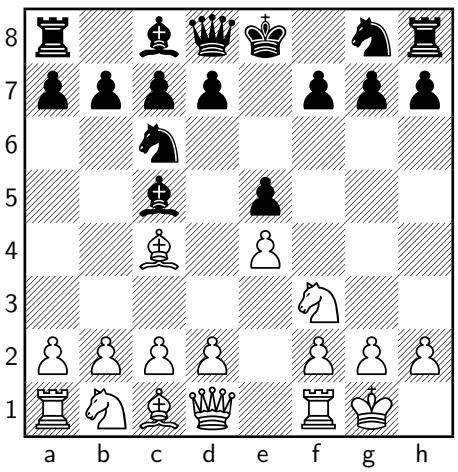
The number of unique pieces moved has almost no influence on the outcome of a game, most likely because both players in any one game already move a high number of unique pieces as we discovered during the feature extraction process. While this feature has a straightforward interpretation, the other features require a much more extensive examination:

Castling

Castling was by far the most influential feature when predicting success in a chess game that we measured.

At the lower levels of play, castling early in the game has strong influence on the outcome of a game compared to the highest level. At the 0-999 rating range, a player castling early likely shows he has some knowledge of chess principles. As opposed to other principles such as center square control, castling takes little tactical knowledge. This means players tend to have little resistance from their opponent and weak tactical players are still able to castle early in many early-game positions. Players, white or black alike, simply need to move a pawn, bishop, and a knight before castling.

1 e4 e5 2 ♘f3 ♘c6 3 ♗c4 ♗c5 4 O-O



An example of white castling as early as possible.

Interestingly, the usefulness of castling early on seems to diminish from an average level of play of 1000-1499 ranking up through the highest level. These diminishing returns may be a result of the complexity of the high-level games. Players at the top level cannot simply castle early and hope for an advantage. While both average and low-level players benefit greatly from utilizing castling, at higher levels, players are capable of more complex strategies that go above any advantage these principles provide.

Overall, castling was undoubtedly the best predictor for who wins the chess game. This gives credence to the idea that king safety, as well as activating the rook are very important for most chess players.

Center Square Control

A player's center square control was the second most influential feature when predicting if a player won a chess game.

In the lowest rating bracket, the impact of this feature is nearly double that in any other rating range. We theorize that since players at the lowest rating range have a more primitive understanding of chess strategy, having pieces at the center provides a natural advantage. When players have pieces in the center, they have more control of the

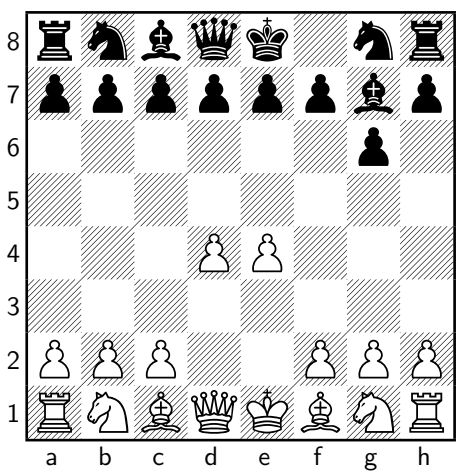
board and therefore more tactical opportunities. Players with more tactical opportunities are at a significant advantage because they are more easily able to launch tactical combinations on their opponents.

While the number of the center squares controlled proves to be important in low level games, its does not have as significant of an impact in the majority of mid and upper level games. This is quite interesting as most players would consider have control of the center board as an advantage. Our results show that while it is a beginner friendly strategy that can help early on, once a player attains a better understanding of the game, other strategies become more important.

We were surprised to see that center square control still had influence on the outcome of games in the highest rating bracket. A popular opening style called "hypermodern" openings are widely used in this rating range with much success. These openings completely give up control of the center early on which breaks this opening principle. These setups instead focus on creating a strong attack and "break" open the center later on.

These results show the great understanding of the game that higher ranked individuals have.

1 e4 g6 2 d4 ♗g7



An example of black giving up center control in a hypermodern setup.

Gambits and Unique Pieces Moved

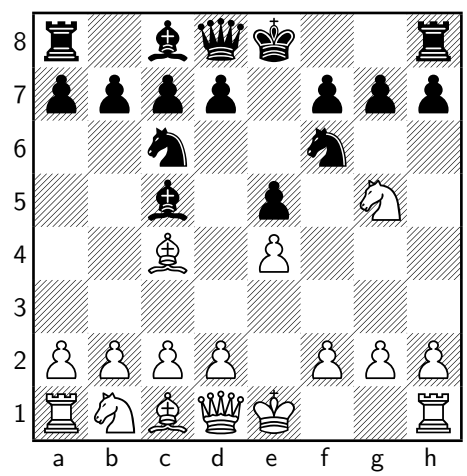
Neither whether a gambit was played or not, nor the number of unique pieces moved could be used to predict which player wins in any skill bracket.

We hypothesised that gambits would be most successful in 1500-1999 category. This prediction was correct, although gambits as a whole had a very low impact on the outcome of a game.

We were surprised to see that playing a gambit opening was such a poor predictor of a game's outcome across all rating ranges. This was surprising because we expected that playing such aggressive openings would be very successful in the mid-level ratings and strongly correlate with wins. It is likely that gambits tend to complicate the game too much, leading to the abandonment of opening principles. Once complicated positions arise in a game, most chess opening principles are ignored and careful calculations take their place.

The number of unique pieces moved early in the game also had little influence when predicting who won a chess game. We believe this is because moving unique pieces is such a common strategy. Since so many people use this strategy, it is not very predictive of the game's outcome.

1 e4 e5 2 ♘f3 ♘c6 3 ♗c4 ♘f6 4 ♘g5 ♗c5



An example of a tactical position where principles must temporarily be abandoned.

VII. Conclusion

Our goal was to discover the success of various chess opening principles in different rating brackets. Of the principles selected, we found that the turn when a player castles held the highest influence across all levels of play. We also found that as the skill of players increases, opening principles become less relevant as more complex strategies arise.

The biggest take away seems to be that many chess principles act only as guidelines and don't always equate to a greater advantage, especially as the level of players increases. This is shown most apparently in the highest rating range, where every chess principle has very little influence on game outcome. Chess is simply that complex of a game.

(But make sure you castle)

VIII. References

[1] Sadikov A., Možina M., Guid M., Krivec J., Bratko I. (2007) Automated Chess Tutor. In: van den Herik H.J., Ciancarini P., Donkers H.H.L.M.. (eds) Computers and Games. CG 2006. Lecture Notes in Computer Science, vol 4630. Springer, Berlin, Heidelberg.

[2] The Board of the ICCA: The Best Annotation Award for 1993. ICCA Journal 17(2), 106–108 (1994)

[3] Chessmaster at Metacritic. Metacritic.