Chess and Individual Differences

Angel Blanch

CHESS AND INDIVIDUAL DIFFERENCES

Research from the neurosciences and behavioural sciences highlights the importance of individual differences in explaining human behaviour. Individual differences in core psychological constructs, such as intelligence or personality, account for meaningful variations in a vast range of responses and behaviours. Aspects of chess have been increasingly used in the past to evaluate a myriad of psychological theories, and several of these studies consider individual differences to be key constructs in their respective fields. This book summarizes the research surrounding the psychology of chess from an individual- differences perspective. The findings accumulated from nearly forty years' worth of research about chess and individual differences are brought together to show what is known – and still unknown – about the psychology of chess, with an emphasis on how people differ from one another.

Angel Blanch works in the Department of Psychology at the University of Lleida, Catalonia, Spain. His research focuses on individual differences, intellectual performance and data analyses in behavioural science. Angel also serves as an associate editor at *Personality and Individual Differences*, as well as advising on the editorial board for *Psychological Assessment and Stress and Health*.

CHESS AND INDIVIDUAL DIFFERENCES

ANGEL BLANCH
Universitat de Lleida

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108476041
DOI: 10.1017/9781108567732

First published 2021

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Blanch, Angel, 1967– author.
Title: Chess and individual differences / Angel Blanch, Universitat de Lleida.
Description: New York, NY : Cambridge University Press, 2021. | Includes bibliographical
references and index.
Identifiers: LCCN 2020031750 | ISBN 9781108476041 (hardback) | ISBN 9781108567732 (ebook)
Subjects: LCSH: Chess – Psychological aspects. | Chess – Social aspects. |
Pattern perception. | Individual differences.
Classification: LCC GV1448 .B53 2021 | DDC 794.1–dc23
LC record available at https://lccn.loc.gov/2020031750

ISBN 978-1-108-47604-1 Hardback

CONTENTS

FIGURES

TABLES

PREFACE

A considerable body of research within several fields of neurosciences and behavioural sciences has highlighted the crucial importance of individual differences in explaining human behaviour. Individual differences in core psychological constructs such as intelligence or personality account for meaningful variations in a vast diversity of responses and behaviours. Some aspects of the game of chess have been used in the past to evaluate a myriad of psychological theories. Several of these studies consider individual differences as key constructs in their respective fields of research. This book summarizes the latest research about the psychology of chess from an individual differences approach. The volume provides a comprehensive overview of the findings accumulated through nearly forty years of research into chess and individual differences. This volume, *Chess and Individual Differences*, organizes a complete perspective in terms of what is already known and what remains unknown about the psychology of chess, with an emphasis on individual differences.

ACKNOWLEDGEMENTS

Writing this book would have been impossible without the help of the Cambridge staff. In particular, I am gratefully indebted to Janka Romero, Emily Watton, and Jessica Norman for providing assistance throughout the process. Special thanks go to Guillermo Campitelli, who elaborated extensive and priceless feedback on an earlier draft of the manuscript. My greatest debt, however, is to my wife, Loles, and to my two-year-old daughter, Petra, who stoically withstood the time and effort spent on the book and gave me comfort in the meanwhile.

1

Introduction

Several facets of the game of chess have been used in the past to model and evaluate a myriad of psychological theories in a variety of empirical studies. Most of these studies have taken either an experimental or a correlational approach (Table 1.1). Over half a century ago Lee Cronbach examined in detail the evolution of empirical psychology stemming from these two lines of work (Cronbach, 1957). Cronbach contended that a combination of the experimental and correlational approaches would be the most rewarding for advancing psychology, in both basic and applied research. Analogous arguments have repeatedly been brought up, while advocating for a greater degree of cooperation between cognitive scientists and differential psychologists regarding the study of human intelligence (Deary, 2001). Individual differences in several psychological attributes other than intelligence are critical for understanding the behaviour of people. In the past forty years there has been growing interest in the role of these individual differences, because they appear to modulate human behaviour in important domains such as work, health, and education.

Chess can provide a commensurate model of human behaviour, akin to the *Drosophila* model in the biological sciences (Simon & Chase, 1973). Chess has typically been used in terms of the experimental approach to model several theories concerned with cognitive psychology topics. Moreover, the studies carried out in the domain of chess have also increasingly suggested that there are individual differences in several human behavioural attributes, such as brain functioning, memory, thinking, decision-making, intellectual human performance, personality, and motivation. This book compiles and describes this latter body of research.

1.1 A Very Brief Opening to the Game of Chess

The origins of the game of chess can be traced back to ancient India around the sixth century AD. Chess travelled first to the West, then, later, to the rest of the world. Nowadays chess has become the universal intellectual game par excellence, practised by millions of individuals of diverse nationalities, ages, and backgrounds. Chess is played on an eight by eight squared board, divided into thirty-two light squares and thirty-two dark squares. Each square is uniquely

Table 1.1 *Overview of the two main approaches to psychological research*

	Experimental	Correlational
Aim	Functional analyses of psychological processes	Analysis of individual differences and regularities in behaviour
Unit of analysis	Cognitive processes	Psychological traits
Hypotheses	Inference	Covariation
Research design	Experimental	Ex-post-facto Probabilistic
Data analyses	ANOVA ANCOVA MANOVA	Correlation Factor analysis Causal analyses
Validity	Internal	External

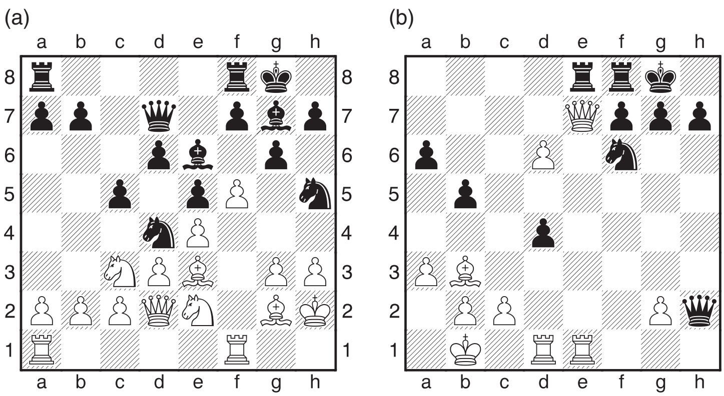

Figure 1.1 A chess game with all intervening pieces in action (left diagram); a chess problem with white to play and win (right diagram; taken from a game between Velmirovic and Csom in Amsterdam, 1974)

identified by a coordinate system using Latin letters from a to h and numbers from 1 to 8. This board imitates a battlefield on which two armies, one black and one white, confront each other in a merciless fight. Each of the two armies comprises eight pawns, two rooks (R), two knights (N), two bishops (B), a queen (Q), and a king (K). The left diagram in Figure 1.1 shows an ongoing typical chess clash with all these intervening pieces. The specific moves of all pieces are described briefly in the Glossary, together with the value of each piece, indicated by the points usually assigned to it. The aim of the game consists in checkmating the opponent's king. The army that first checkmates the enemy king wins.

The basic rules of the game are very simple and very easy to learn, even at younger ages and at any educational level. Yet the game as a whole becomes extremely complex. There are literally several millions of millions of different combinations among the contending pieces in a single chess game. These combinations can be represented with a chess tree, an informational device in which the solution is the path leading to victory. A chess tree typically generates a massive and unmanageable amount of combinations (10^{120}) even for the most powerful and fastest computer chess engines, let alone for human beings (Shannon, 1950). Because each of the pieces involved in the game obeys different movements, the game is intellectually demanding, while requiring the interplay of a variety of major psychological attributes and processes, such as perception, memory, reasoning, decision-making, problem solving, will, motivation, interests, and creativity.

Consider, for instance, the right diagram in Figure 1.1. This represents a typical chess problem with the white forces to play and win. There is an efficient sequence leading to the white victory that comprises five precise moves, with an average time limit to solve it of about ten minutes. The correct sequence of moves in algebraic notation is shown below, for both white and black pieces:

	White	**Black**
1	B×f7+	R×f7
2	Q×e8+	N×e8
3	R×e8+	Rf8
4	d7	Qd6
5	Rf1!!	1–0

Each of the five chess moves comprise two plies: one ply for white, and one ply for black. The ply corresponding to black in the fifth move indicates that white has won the game (scoring one point), however, because, after the last ply of white (Rf1), there is no possible legal move by black to avoid being checkmated at the very next move by white. Capital letters stand for the specific chess piece being moved and the × symbol indicates that a piece captures an opponent's piece. For instance, the first ply for white (B×f7+) indicates that the white bishop (B), initially placed in the b3 square, is capturing the pawn located in the f7 square. The + symbol indicates that the black king is placed in check. A ply depicting a single square only indicates a pawn move. For instance, the fourth ply for white (d7) indicates that the pawn placed in the square d6 advances to the square d7. The double exclamation mark in the fifth ply for white (Rf1!!) indicates a brilliant and very strong move. In this specific game, it was a *coup de grâce* move, winning the game.

People may differ greatly in terms of their chances of finding out this sequence of moves. If you are a proficient chess player at the master level, you may be able to 'see' the sequence at a glance. It could also be the case that you may remember this position because you have already studied it in the past, during your long chess career. On the other hand, if you are a typical club chess player with a moderate level of chess skill, you might invest the suggested amount of time, but you may end up unable to figure out what the correct solution is at all. If you are a beginner chess player, or you just know the basic chess rules, the likelihood of experiencing serious difficulties in finding the solution may be so great as to be insurmountable. This is a very basic example of individual differences in chess performance and chess skill.

1.2 Overview of This Book

Nowadays there is a considerable volume of chess studies that have highlighted noteworthy individual differences. For example, some of these chess studies use problems such as that shown in Figure 1.1 as experimental stimuli. This book is an attempt to compile and summarize the latest research about the psychology of chess with a focus on individual differences. Besides, this volume aims to provide an overview of the findings from more than forty years of research, from the mid-1970s to date, about chess and individual differences. This body of research has sometimes yielded inconclusive and even controversial results, suggesting, for instance, that the development of chess skill over time may largely depend on the combination of individual differences in several traits or broad clusters of traits. This book organizes the body of knowledge that uses chess as a model environment, while providing useful scientific information about a variety of individual differences in brain functioning, intelligence, personality, expertise, and sex, and in applied fields such as business, health, and education.

The book is mainly aimed at scholars within the broad spectrum of the social and behavioural sciences who have an interest in the psychology of chess. The book can be of interest to psychologists, sociologists, educators, neuroscientists, and behavioural scientists in general. The chapters are intended to cover the topics typically addressed by social scientists interested in individual differences working in a diversity of fields. Those researchers and academics working in brain functioning, human abilities, and personality may find the book appealing. Moreover, the book may also arouse the curiosity of researchers and academics working with topics such as expertise, sex differences, and education, or with a focus on applied fields. In addition, the book may also be of interest for people who play chess themselves. In particular, chess players wishing to gain a more in-depth understanding of the scientific work undertaken with chess as a model domain from a psychological approach may find some stimulating information within these pages.

Chapter 2 describes the Elo chess rating. What makes chess an optimum field for the study of individual differences is the availability of this objective quantitative measure to gauge a player's chess strength. The Elo rating system is by far the most popular and accepted indicator worldwide for quantifying accurately individual differences in chess skill. Every chess player participating regularly in rated chess tournaments holds an Elo rating. The Elo rating changes according to the outcomes of the games played within a given time period, while considering the Elo rating of the opponents. The chapter describes how the Elo ratings of thousands of chess players are kept and periodically updated. It also outlines the updating mechanisms and some basic statistics of the Elo rating. In addition, the chapter describes some recent alternatives to rating chess skill, such as the Universal Rating System (URS). Appendix 1 summarizes the studies that have used the Elo rating as related to a variety of human behaviours.

Chapters 3 and 4 provide an overview of the main findings from the cognitive and the individual differences approach to the psychology of chess, respectively. Chapter 3 reviews the main research findings from the cognitive or experimental paradigm within psychology, which originated with the precursor scientific works about the psychology of chess. Three main basic facets of human behaviour have been addressed within this general approach: perception, memory, and thinking. The main conclusions from this extensive body of research can be summarized by emphasizing the role of individual differences. Chapter 4 outlines the main tenets and constructs of differential psychology, the discipline that studies individual differences in behaviour relevant for central social realms such as health, education, and work. The chapter is structured around three main themes. First, it describes the characterization and appraisal of individual differences. Second, the PPIK theory is suggested as an optimal starting point to conceptualize and examine individual differences. This framework comprises traits from four broad dimensions: intelligence as process, personality, interests, and intelligence as knowledge. Third, the chapter closes by addressing the old but compelling debate about the heredity versus environment dichotomy in explaining complex human intellectual behaviour.

Chapter 5 describes the studies addressing human biological factors in chess, with a focus on psychophysiology and brain imaging. Human psychophysiology is a multi-faceted and complex phenomenon. The game of chess has provided a proper domain for the study of the central psychophysiological mechanisms underlying psychological processes such as stress, emotion evaluation, and decision-making. Moreover, novel technologies designed to provide high-resolution brain imaging are being increasingly used to explain human behaviour. These technologies have also been used with chess players to examine the interrelationships of brain and cognitive functioning, and with personality and intelligence factors. In particular, this chapter outlines the research undertaken with electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and positron emission tomography (PET). The chapter summarizes this body of evidence while underlining the most significant conclusions that may be derived from this intriguing and thought-provoking field of research.

Chapter 6 provides an account of the studies addressing chess and intelligence. Human intelligence is one of the main general objects of study in individual differences research. There are indeed multiple models about and approaches to human intelligence, which are briefly described within this chapter. Chess has been typically associated with a high level of intelligence. Whether chess players are more intelligent on average than the general population is a recurrent question that has elicited a considerable body of research. There are unsettled issues as to what constitute the most advantageous cognitive abilities required in chess, and whether playing chess makes people smarter. These topics have been addressed with both children and adults. The scientific evidence in connection with this topic is inconclusive, however, and controversial in some instances. This chapter addresses these matters of contention by summarizing the state of the art in this particularly cogent field of research. The final section in the chapter includes novel empirical findings comparing chess skill and chess motivation in the prediction of chess performance, suggesting that non-cognitive traits might also be influential for chess performance.

Chapter 7 analyses what is already known about chess and human personality. Personality is the other main broad domain addressed within the general framework of individual differences. In contrast with intelligence, however, the body of research concerning the personality of chess players is rather scarce. There have been some interesting findings recently, however, and these are summarized within this chapter. After describing briefly the main approaches to addressing human personality, some questions addressed in this chapter are whether personality influences chess playing style, or whether a chess player's personality differs in some special way from that of other people. In addition, whether personality factors may interact with cognitive abilities in chess players is an interesting and relatively novel topic. The chapter closes by presenting novel data about the interplay between personality, motivation, and emotional regulation in predicting chess skill.

Chapter 8 analyses expertise, one of the most prolific fields in empirical research using chess as a model domain. Expertise is of great importance in several realms of human intellectual activity. The role of practice in the development of chess expertise is reviewed in detail in this chapter. Moreover, the role of practice is contrasted with talent, because the deliberate practice approach has advanced the idea that expert performance depends exclusively on practice. A consistent body of evidence suggests that deliberate practice alone is unable to explain the individual variability in chess expertise, however. The present chapter addresses this controversy by framing these findings in the nature versus nurture debate, one of the central themes within individual differences research. Furthermore, this chapter also explores age-related cognitive decline in human intellectual activity, which appears to occur to a lesser extent in the chess domain. For instance, recent findings suggest in particular two interrelated factors that may be highly relevant in preventing cognitive decline in chess: the level of expertise attained, and the amount of tournament activity.

Chapter 9 tackles the issue of sex differences in chess. On average, men tend to start earlier, perform at a higher level, and persist longer than women in the chess domain. Moreover, women are highly underrepresented in chess, which is also apparent in several other domains, such as those connected with STEM fields (science, technology, engineering, and mathematics). The marked difference in the number of men and women participating in chess has led to the assumption that the differences in chess performance between men and women are attributable to a statistical effect derived from the differences in participation rates. In contrast, other findings suggest that men might have an innate advantage in terms of chess playing, enhanced by certain cultural factors. These two points of view are addressed in this chapter. The alternative explanation to the marked disparity in chess participation and performance between the two sexes may be related to the participation of men and women in STEM fields. In addition, there are some noticeable differences in the chess playing of men and women, even though women are able to play very strong chess, just like men. The chapter closes by presenting a statistical analysis with data from the chess domain, which relates to sex differences in performance at different levels of practice. The findings from this analysis suggest that sex differences in the Elo ratings tend to increase with increasing practice, pointing to factors other than practice as the underlying causes of these sex differences.

Chapter 10 deals with the applications of chess in three major fields of human activity: business, health, and education. Chess has been used in the business field with two main aims. First, chess has been used for educational purposes to teach and consolidate concepts connected with this discipline. Second, some studies have used chess as a model to evaluate game-theory aspects of the game. The game of chess has also been increasingly used to address health-related problems such as attention deficit hyperactivity disorder (ADHD), neurodegenerative disorders, and schizophrenia. Moreover, chess has become an increasingly popular pedagogical method in several school settings across the world. A number of studies claim that chess training entails several educational benefits for core academic subjects such as languages and mathematics, and also for concentration and self-control, or the development of socio-affective competences. Several of the instructional experiences that use chess to enhance these behaviours are described in this chapter. Some recent studies suggest that significantly higher levels of academic performance for schoolchildren and adolescents are associated with chess-based teaching or the practice of chess on a regular basis, when compared with those students who are not involved in chess playing or chess instruction. Another set of studies have questioned the purported benefits of chess training for formal education, however. From this latter point of view, there are both conceptual and methodological concerns that compromise to a great extent the available evidence about the association of chess training with academic achievement. Two of these issues relate to the transfer of abilities across domains, and to the concept of statistical power.

Chapter 11 is the closing chapter of this book. This chapter argues why chess has become an interesting domain to address topics of interest for individual differences research. It also summarizes the most robust available evidence to date by outlining the key findings, while suggesting some tentative and potentially promising steps for advancing the field.

2

Quantifying Chess Skill

What makes chess an optimum field for the study of individual differences is the availability of an objective quantitative measure of a player's chess strength. This is an important asset compared with other applied domains, because they lack such a systematic indicator of skill. Although several indicators quantify accurately chess skill, the Elo rating system is the more popular chess skill indicator, accepted worldwide. For example, the Elo rating system is useful in the organization of formal chess tournaments, such as in pairing players of equivalent chess strength, or in restricting participation in chess tournaments to a given chess strength level or to groupings of players with different levels of chess skill. Because the Elo rating is an interval scale, it lacks a true zero, though it allows the quantification of an objective difference between each value.

Every chess player participating regularly in rated tournaments holds an Elo rating. The Elo rating is a dynamic indicator that depends on the outcomes of the games played within a given time period, taking into account the Elo rating of the opponents. Such a system has been deemed highly appropriate to track changes in the variability of its scale values, which might be useful for addressing an extensive variety of research problems within differential psychology or individual differences research (Batchelder & Bershad, 1979; Howard, 2006).

A sense of the variability in chess skill as measured by the Elo rating can be gleaned by looking at the world maps displayed within Appendix 2. These maps represent data for 118 countries obtained from the December 2018 list of the World Chess Federation (Fédération Internationale des Échecs, FIDE). The first map shows the mean Elo rating, and the second map shows the number of chess grandmasters by country. There are only three countries with a mean Elo rating above 2700 Elo points – Russia, China, and the United States – and twelve countries with a mean Elo rating above 2600 Elo points: Azerbaijan, Ukraine, India, France, Armenia, Hungary, the Netherlands, Poland, the United Kingdom, Israel, Germany, and Spain. Cross-country differences are more pronounced, however, when looking at the number of grandmasters in the second map. Here, Russia holds a noteworthy advantage over the rest of the countries, with 251 grandmasters, in front of the United States, with ninety-eight, Germany, with ninety-six, and Ukraine, with ninety-one. There is then a group of five countries with between fifty and fifty-seven grandmasters: Serbia, Hungary, India, Spain, and France. In contrast, the world regions with the lower mean Elo ratings and number of grandmasters correspond to Africa, and several countries in Central and South America, and Asia. The two maps evidence the universality of the game, which is surely unparalleled by any other game of its kind.

2.1 Elo Rating Lists

The systematic updated records of the Elo ratings of chess players from all over the world allow the study of individual differences in intellectual performance from an objective point of view. Every chess player participating regularly in rated chess tournaments of any kind holds an Elo rating that ranges from approximately 1,200 to about 2,850 points, with higher scores being indicative of a higher level of chess strength (Elo, 1978; Glickman, 1995; Glickman & Chabris, 1996; Glickman & Jones, 1999). Chess federations worldwide keep and update periodic records of the Elo ratings of their respective players. In addition, players participating in international tournaments hold the Elo rating of the respective player's country or local chess federation, and the international Elo rating assigned by the World Chess Federation (Fédération Internationale des Échecs: FIDE). Elo ratings from different chess federations tend to be highly correlated. There are even Elo rating lists from a variety of computer chess engines. Figure 2.1 shows part of the Elo rating lists of the FIDE, the Spanish Chess Federation, and the Catalan Chess Federation, and the Elo ratings of 353 computer chess engines.

The lists from the World and Catalan Chess Federations and from computer engines are ordered by the rank of the strongest players. The World Chess Federation list shows the ten strongest players. The Catalan Chess Federation list indicates the chess title and sex of each player (GM: Grandmaster; M: Male). The computer list shows the number of games played and the percentage of winning outcomes. The Spanish Chess Federation list is in alphabetical order by the player's surname, and it also includes the number of games played in the given period, the year of birth, the title, whether the player is active or inactive (A, I), and the previous Elo rating. For instance, the current Elo of the first player in this list is 1851 points, while his previous Elo was 1842. Therefore, the player has gained nine Elo points in this latter Elo update. In contrast, the player with Id. FEDA #26 has a current Elo of 2177, while his previous Elo was 2181. Therefore, this player has lost four Elo points in this latter Elo update.

(a)

World Chess Federation

Rank	Name	Title	Country	Rating
1	Carlsen, Magnus	g	NOR	2853
2	Anand, Viswanathan	g	IND	2816
3	Topalov, Veselin	g	BUL	2816
4	Nakamura, Hikaru	g	USA	2814
5	Caruana, Fabiano	g	USA	2797
6	Giri, Anish	g	NED	2791
7	Kramnik, Vladimir	g	RUS	2783
8	So, Wesley	g	USA	2780
9	Grischuk, Alexander	g	RUS	2771
10	Aronian, Levon	g	ARM	2765

(b)

Listado de Jugadores FEDA

Febrero 2016

Id. FEDA	Nombre del Ju	Federación	ELO	Partidas	Año Nac.	Titulo	Act/Inact	Elo Anterior
25303	ALSHAMEARY F	AND	1851	9	1997		A	1842
21324	ALSHAMEARY F	AND	2315	9	1992	MF	A	2299
4212	ALSINA FRANC	ARA	1875	0	1987		A	1875
14229	ALSINA KIRCHN	CAT	1929	0	1958		I	1929
13807	ALSINA LEAL, D	CAT	2514	0	1988	GM	A	2514
15380	ALSINA LOPEZ,	MEL	1837	3	1993		A	1794
24709	ALSINA LOPEZ,	MEL	1397	0	1996		A	1397
4213	ALSINA MARTI,	CAT	1845	0	1978		I	1845
26	ALSO ALVARAD	CAN	2177	1	1969		A	2181
29754	ALTABAS FELIP	CAT	1592	0	1981		A	1592
4214	ALTAFULLA SA	CAT	1845	0	1942		A	1845

(c)

Federació Catalana d'Escacs

1	codi	nom	elo	titol	sexe
2	23763	NAVARA, DAVID	2710	GM	M
3	27241	VALLEJO PONS, FRANC	2696	GM	M
4	5953	ILLESCAS CORDOBA, M	2615	GM	M
5	17068	PERALTA, FERNANDO	2610	GM	M
6	21349	LAZNICKA, VIKTOR	2607	GM	M
7	23575	FIER, ALEXANDRE	2605	GM	M
8	17936	KOGAN, ARTHUR	2601	GM	M
9	21893	AROSHIDZE, LEVAN	2600	GM	M
10	7459	LOPEZ MARTINEZ, JOS	2595	GM	M

(d)

```
THE SSDF RATING LIST 2016-08-13   %138040 games played by  353 computers

                                            Rating   +     -  Games   Won
                                            ------  ---   --- -----   ---
  1 Komodo 9.1 MP x64 2GB Q6600 2,4 GHz      3366   30   -27   987   82%
  2 Stockfish 6 MP x64 2GB Q6600 2,4 GHz     3335   31   -28   776   77%
  3 Komodo 7.0 MP x64 2GB Q6600 2,4 GHz      3276   31   -29   694   74%
  4 Komodo 5.1 MP x64 2GB Q6600 2,4 GHz      3243   24   -23   956   66%
  5 Stockfish 3 MP x64 2GB Q6600 2,4 GHz     3207   22   -21  1170   67%
  6 Deep Rybka 4 x64  2GB Q6600 2,4 GHz      3202   22   -21  1193   69%
  7 Deep Rybka 3 x64  2GB Q6600 2,4 GHz      3193   22   -21  1371   75%
  8 Deep Hiarcs 14 2GB Q6600 2,4 GHz         3191   21   -20  1210   65%
  9 Chiron 3.01 MP x64 2GB Q6600 2,4 GHz     3166   56   -61   149   38%
 10 Naum 4.2 MP x64 2GB Q6600 2,4 GHz        3145   21   -21  1083   61%
 11 Deep Junior Yokoh x64 2GB Q6600 2,4 GHz 3123   26   -26   690   50%
 12 Naum 4 x64  2GB Q6600 2,4 GHz            3118   21   -20  1276   66%
 13 Deep Junior 13.3 x64 2GB Q6600 2,4 GHz  3117   21   -21  1088   53%
 14 Hiarcs 14 256MB Athlon 1200 MHz          3105   34   -33   440   62%
 15 Deep Shredder 12 x64 2GB Q6600 2,4 GHz  3103   18   -18  1507   62%
 16 Spike 1.4 MP 2GB Q6600  2,4 GHz          3103   18   -18  1509   59%
 17 Deep Hiarcs 13.2  2GB Q6600 2,4 GHz      3100   26   -25   752   59%
 18 Hiarcs 13.1  2GB Q6600 2,4 GHz           3099   24   -24   828   58%
 19 Deep Fritz 13 2GB Q6600 2,4 GHz          3097   24   -24   826   55%
 20 Deep Fritz 12 2GB Q6600 2,4 GHz          3090   20   -20  1200   55%
 21 Deep Rybka 3  256MB Athlon 1200 MHz      3073   39   -37   332   58%
 22 Deep Junior 12 x64 2GB  Q6600 2,4 GHz    3071   20   -20  1178   57%
 23 Deep Fritz 11  2GB Q6600 2,4 GHz         3059   18   -18  1504   62%
 24 Zappa Mexico II x64  2GB Q6600 2,4 GHz   3057   23   -23   900   57%
 25 Naum 3.1 x64  2GB Q6600 2,4 GHz          3045   23   -23   858   53%
 26 Crafty 25.0 MP x64 2GB Q6600 2,4 GHz     3022   26   -27   724   37%
 27 Deep Hiarcs 12  2GB Q6600 2,4 GHz        3017   18   -18  1507   52%
 28 Deep Shredder 11 x64 2GB Q6600  2,4 GHz 3013   21   -21  1088   50%
```

Figure 2.1 Elo rating lists of the World (a), Spanish (b), and Catalan Chess Federations (c), and Elo rating list of top twenty-eight computer chess engines out of a list of 353 engines (d)

2.2 Updating Mechanism and Basic Statistics of the Elo Rating

The Elo rating is a dynamic indicator that depends on the outcomes of the games played within a given period considering also the Elo ratings of the opponents. To illustrate, the elements implied in the calculation and updating of the Elo rating after the outcome of a chess game between two players can be described in three main steps (Elo, 1978; Glickman, 1995):

1. There are three possible outcomes arising from a chess game: win = 1, draw = 0.5, defeat = 0.
2. The expected score (E) in a chess game between players A and B with ratings R_A and R_B can be calculated for player A with the expression in Equation 2.1:

$$E = \frac{10^{R_A/400}}{10^{R_A/400} + 10^{R_B/400}} \tag{1}$$

3. The update of the Elo rating is calculated with the expression in Equation 2.2. This includes a previous Elo rating (r_{pre}), the Elo rating after a chess tournament (r_{post}), a constant value (K), the sum of points obtained in the tournament (S), and the sum of expected scores in each game (S_{exp}):

$$r_{post} = r_{pre} + K(S - S_{exp}) \tag{2}$$

In a single chess game, a win scores one point, a defeat scores zero points, and a draw scores half a point. The expression in Equation 2.1 can be conceived as the actual probability of winning a chess game considering the Elo rating of both opponents. The Elo rating is indeed an accurate predictor of the outcome of a chess game. A stronger player in terms of a higher Elo rating has increased chances of scoring one point when playing against a weaker player. In contrast, a weaker player in terms of a lower Elo rating sees his or her chances of scoring one point greatly decreased when playing against a stronger player. The expression in Equation 2.2 serves to update the Elo rating in accordance with the performance of a player within a given period. The new and updated Elo rating (r_{post}) is calculated by summing the observed previous Elo rating (r_{pre}) and the term $K(S – Sexp)$. The value of K corresponds to an attenuation factor that represents the amount of weight allotted to a new Elo rating given an old Elo rating – that is, the maximum number of points that increase or decrease the rating from the outcome of a single chess game. Larger K values allow greater changes in Elo ratings. Usually, younger and less experienced players tend to have higher attenuation K values than older and more experienced players. The value ($S – Sexp$) indicates the discrepancy observed between the actual points (S) obtained within a given period and the expected points ($Sexp$), calculated with the expression in Equation 2.1 in accordance with the Elo ratings of the corresponding opponents in the chess games played within this period. Positive values in ($S – Sexp$) indicate that the player performed above what was expected, whereas negative values indicate that the player

Table 2.1 *Example of the Elo rating updating in one chess game between players KS versus AB, and JP versus LQ*

Player	Elo rating	*K*	Expected score	Actual score	Game outcome	Elo update
KS	2544	10	0.97	1	Win	2544
AB	1936	15	0.03	0	Defeat	1935
KS	2544	10	0.97	0	Defeat	2534
AB	1936	15	0.03	1	Win	1951
KS	2544	10	0.97	0.5	Draw	2539
AB	1936	15	0.03	0.5	Draw	1943
JP	2064	15	0.44	1	Win	2071
LQ	2106	15	0.56	0	Defeat	2098
JP	2064	15	0.44	0	Defeat	2056
LQ	2106	15	0.56	1	Win	2113
JP	2064	15	0.44	0.5	Draw	2065
LQ	2106	15	0.56	0.5	Draw	2105

performed below what was expected. Therefore, the updated Elo rating will increase or decrease accordingly.

Table 2.1 shows two examples of the Elo rating update in a hypothetical chess game. In the first example, concerning players KS and AB, there is a considerable difference between the Elo ratings of the two players, 608 points, while player KS holds a lower K value of ten compared with the K value of fifteen for player AB. The first two rows in the table show the most likely situation after the game: a victory of the player with the higher Elo rating. The Elo rating update for this player would be unmodified from the previous Elo rating (Elo update = 2544). On the other hand, the Elo rating for player AB would decrease the previous Elo rating (Elo update = 1935) by just one point. In contrast, with an unlikely defeat of the stronger player, KS, the previous Elo rating decreases by ten points (Elo update = 2534), while for the weaker player, AB, there is an increment of fifteen points (Elo update = 1951). The result of a draw, with 0.5 points for each player, is also more advantageous for the weaker player; it decreases the Elo rating of KS by five points (Elo update = 2539), while increasing the Elo rating of AB by seven points (Elo update = 1943). In the second example, the difference between players JP and LQ is only forty-two Elo points, which is markedly lower than in the previous example, with both players holding a *K* value of fifteen. The three possible outcomes of the game, win, draw, and defeat, indicate a more balanced outcome for each player concerning their corresponding Elo updates. This is perhaps better seen

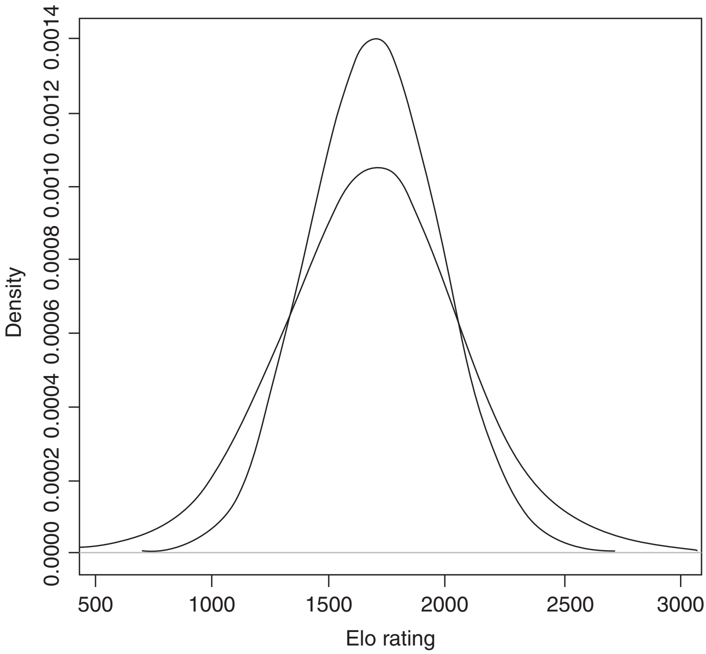

Figure 2.2 Density plot of the Elo rating with normal (continuous line) and logistic (dotted line) distributions

in the last two rows, which describe the draw situation, which modifies the Elo update by just one point for each player.

The Elo rating system can be framed within the Bradley–Terry model for predicting the comparison of a pair of objects or individuals (Bradley & Terry, 1952), albeit, in its development, the Elo rating system assumed a normal distribution. Nevertheless, the Bradley–Terry model relies instead on logistic distribution assumptions, which is the approach taken by the World and the US Chess Federations (Glickman, 1995). Figure 2.2 shows both normal and logistic probability distributions for a simulated sample of 1,000 players with a mean of 1700 Elo points (*Sd* = 200). There are higher density estimates at the centre of the normal distribution, and longer extended tails for the logistic distribution. On the other hand, Figure 2.3 shows the density plots from four different chess tournaments in four variables: the age of the participants, the number of games prior to the tournament, the tournament outcome, and the Elo rating. The data corresponding to the four tournaments indicate a relatively consistent overlap apart from for the Elo rating, suggesting that the Elo rating was more variable than age, number of games, or tournament

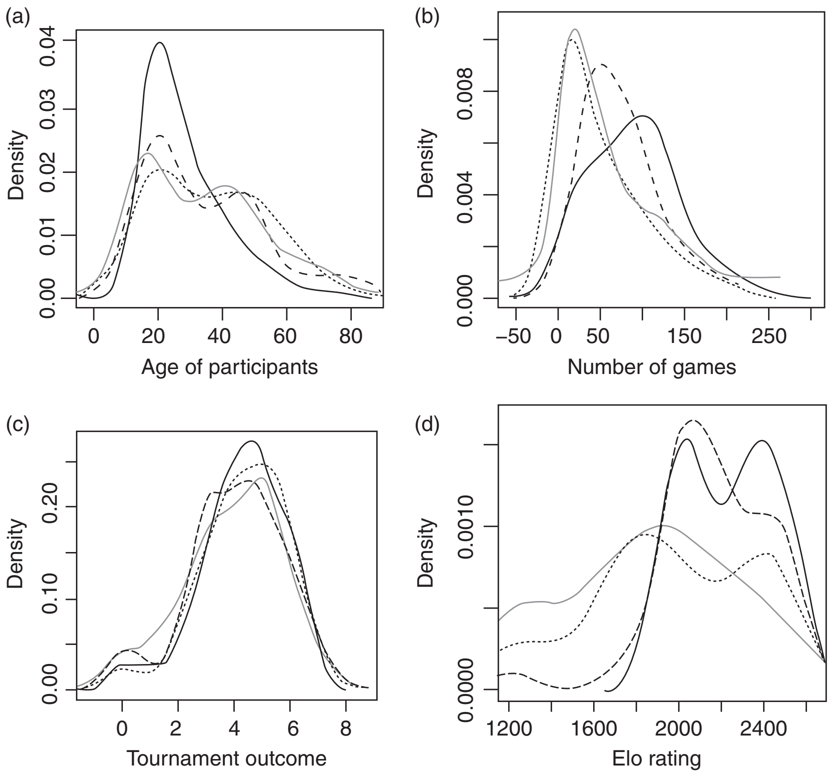

Figure 2.3 Density plots for the distribution of the Elo ratings in the age of participants (a), number of games (b), tournament outcome (c), and Elo rating (d) of four chess tournaments

outcomes. Table 2.2 displays the descriptive statistics in the Elo rating across the four chess tournaments. The Shapiro–Wilk normality tests for the Elo rating in the four tournaments suggest that the null hypothesis stating that the sample comes from a population with a normal distribution is unsupported by these data.

2.3 Alternatives to the Elo Rating of Chess Players

The Elo rating system has gained considerable acceptance and has been used worldwide for over forty years since its inception (Elo, 1978). There are several criticisms and suggestions for improvement, however. For example, it has been recommended that chess ratings should deal with unrated and recently rated players, and that the *K* attenuation factor should be related to the number of

Table 2.2 *Descriptive statistics of the Elo ratings in four chess tournaments*

Statistic	Tournament 1	Tournament 2	Tournament 3	Tournament 4
n	107	100	85	81
Min	1873	1200	1200	1200
Q1	2034	1591	1724	2025
Q2	2234	1878	1982	2131
Q3	2411	2143	2349	2394
Max	2625	2574	2596	2577
M	2234	1847	1975	2165
Sd	213	386	393	262
Cv	0.10	0.21	0.20	0.12
Skewness	0.01	-0.15	-0.32	-1.03
Kurtosis	-1.27	-0.80	-0.73	2.86
W	0.95***	0.96**	0.95**	0.91***

Notes: n = sample size; Min: minimum; Q1 to Q3: quartiles 1 to 3; Max: maximum; M: mean; Sd: standard deviation; Cv: coefficient of variation (Sd/M); W: Shapiro–Wilk normality test (H_0: the variable follows a normal distribution); $^{**}p < 0.01$; $^{***}p < 0.001$.

games played (Fenner, Levene, & Loizou, 2012). There are, in addition, other alternatives to rate chess skill by relying on Bayesian methods. For instance, there is a proposal to estimate chess skill from the moves played in an assortment of chess games, rather than from competitive chess outcomes (Di Fatta, McHaworth, & Regan, 2009). Other work has focused on the modelling of draws and the inference of chess skill from chess team outcomes (Herbrich & Graepel, 2006).

Furthermore, there are other approaches raising substantial modifications to the Elo rating system. For example, Chessmetrics is a comprehensive internet database about the rating of the chess skill of chess masters throughout history. Jeff Sonas, the developer and chief engineer of Chessmetrics, proposed an alternative system to calibrate chess skill with alternative methods to the Elo rating (Sonas, 2002). There were four main suggestions regarding the Sonas system: using a more dynamic *K*-factor; dismissing the Elo table and opting for a simpler linear model; including faster time controls, albeit assigning them a lower importance than slower time controls; and calculating the chess ratings on a monthly basis. Moreover, other alternative chess rating systems include the Glicko (Glickman, 1999), the Glicko-2 (Glickman, 2001), and the Universal Rating System (URS), all of them developed by Professor Mark Glickman, a statistician at Harvard University, together with other researchers.