

AGE TRENDS IN EXPERT CHESS: IS AGING KINDER TO THE INITIALLY MORE
ABLE?

by

Karin Almuhtadi

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Arts
in Clinical Psychology

at

Lakehead University

November 2011

Supervisor: Dr. Michael J. Stones

Second Reader: Dr. Rupert Klein

External Examiner: Dr. Antony Puddephatt

Internal Examiner: Dr. John Jamieson

Acknowledgements

First and foremost, much thanks to Dr. Stones for his guidance with this study. His knowledge was crucial to the completion of this work.

Thanks to Drs. Klein and Puddephatt, for their critical review and feedback. Your presence on my thesis committee is greatly appreciated.

Lastly, to my other 2/3, thank you for your love and unconditional ways throughout this part of my life process.

Table of Contents

Introduction.....	8
1.1 Defining Expert Performance.....	8
1.2 Acquisition and Maintenance of Expertise.....	10
1.3 Is Aging Kinder to the More Expert?.....	15
1.4 FIDE.....	21
1.5 Chess Research.....	24
1.6 Chess Expertise and Aging.....	30
1.7 Strategic Participation in Sport.....	33
1.8 Rationale.....	36
Method.....	37
2.1 Participants.....	37
2.2 Materials.....	38
2.3 Variables.....	39
2.4 Method of Analysis.....	40
2.5 Statistical Analysis.....	42
Results.....	43
3.1 Preliminary Findings.....	43
3.2 Main Findings.....	43
Discussion.....	55
4.1 Summary of the Findings.....	55
4.2 Limitations.....	58
4.3 Implications.....	59

4.4 Conclusions.....	60
References	61

Tables.....	23
Table 1: Rating system in expert chess: Rating ranges and title categories.....	23
Table 2: Comparison of multilevel models: Sequence of model parameters entered.....	44
Table 3: Estimates of fixed effects for the null model (DV = Rating).....	45
Table 4: Estimates of covariance parameters for the null model (DV = Rating).....	45
Table 5: Estimates of fixed effects for model 10 with no interactions (DV = Rating).....	46
Table 6: Estimates of covariance parameters for model 10 with no interactions (DV = Rating).....	48
Table 7: Estimates of fixed effects for model 18 with interactions (DV = Rating).....	48
Table 8: Estimates of covariance parameters for model 18 with interactions (DV = Rating).....	50
Table 9: Estimates of fixed effects for the final model with higher-order interactions (DV = Rating).....	51
Table 10: Estimates of covariance parameters for the final model with higher- order interactions (DV = Rating).....	52

Figures.....47

 Figure 1: A graph of mean ratings by player's entry age.....47

 Figure 2: A collapsed graph of mean change from baseline rating by elapsed age
 for three age groups and rating categories54

Abstract

The “aging is kinder to the more initially able” hypothesis has received varied support across different domains of expertise, but within the chess domain is largely accepted. Some suggest that tournament participation has little effect on the ratings of older chess experts. This has led to conclusions about the importance of solitary practice for older experts as well as provided support for chess experts maintaining their performance ratings well into old age. Twenty-five World Chess Federation (FIDE) rating lists (2006-2010), consisting of 99,608 players and over 1.5 million observations were assessed using multilevel linear modeling (MLM). Findings demonstrated that aging was indeed kinder to the initially more able, but more importantly, that playing more games was associated with higher ratings. This has implications for aging expert players, such that active tournament participation counts.

Keywords: chess, expertise, aging, MLM

Age Trends in Expert Chess: Is Aging Kinder to the Initially More Able?

Many studies have explored the changes in cognitive and physical functioning with respect to human aging. In order to keep the mind and body functioning well into old age individuals have to take an active part in exercising it. The phrase “use it or lose it” captures this idea well. While we are all bound to *lose it* at some point, the way the aging process unfolds varies significantly between persons. Moreover, researchers have focused on expert performers who remain active well into old age. This group provides researchers with a gold standard of what is possible in a certain expert domain at different points over the lifespan.

A brief description of expert performance precedes proposed ways of acquiring and maintaining expertise. Studies that address the “aging is kinder” hypothesis are then reviewed. A summary of the chess domain, as well as the existing chess literature, follows prior to the topic of chess experts and aging. Finally, before reporting on the proposed objectives and current study, a section is devoted to strategies of participating in sport. A discussion about why older experts may strategically participate in chess wraps up the introduction.

1.1 Defining Expert Performance

The subject of expertise or expert performance has been the focus of research in various disciplines throughout history. Expertise can be exhibited in an endless number of domains. In some research, domains have been roughly described as more physical or cognitive, such as swimming and chess, respectively. The definition of expertise in different domains varies, but the question of “who is the expert performer?” seems to be

of continuing interest to both researchers and the lay public. Ericsson and Lehmann (1996) described expert performance as follows, “constantly superior performance on a set of tasks relevant for a specific skill or topic, which is achieved by deliberate practice over a period of at least 10 years” (p. 277).

In the literature, the following terms are used interchangeably to refer to experts: expert performer, elite, Master, and initially more able. Participants are also referred to as more or less expert. Interestingly, researchers found that two works that addressed chess experts (de Groot, 1946/1978; Simon & Chase, 1973) were often inaccurately cited (Bilalić, McLeod, & Gobet, 2008). The authors point out that participants in de Groot’s study were all experts, yet a number of articles they reviewed referred to them as experts and novices. Imprecise and inconsistent use of terminology can be problematic for comparison of studies.

Chi (2006) describes two broad ways of studying expert performance. These include studying present day experts retrospectively and comparing experts to non-experts. Yet another way of studying expert performance is comparisons within the expert group. This latter design has been less common, with most research comparing experts to novices (Roring & Charness, 2007). Chi also recommends that ways in which experts fall short are an important area of study. The next section outlines one systematic approach to the study of expertise.

Ericsson and Charness (1994) conducted studies of experts in various domains, from which they developed the reproducible superior performance model. Ericsson, Roring, and Nandagopal (2007) further expand on this framework. These researchers suggested the following criteria for the study of expertise: phenomena must be observable

and correspond to measurable performance, associated performance must be generated under controlled and standardized conditions, associated performance can be elicited repeatedly, and performance on the representative tasks must be reproducibly superior to motivated control groups with different amounts of experience in the task domain.

Additionally, one works backwards from identified reproducibly superior performance to trace the development of possible mediating mechanisms. This model places a strong emphasis on empiricism, unlike the anecdotal or subjective evidence that often supported early expert performance research.

1.2 Acquisition and Maintenance of Expertise

How expertise develops over time has been of major interest. The whole life cycle of an expert can be broken down developmentally into a long phase of acquisition of the superior skill, followed by peak and maintenance period, and inevitably, some form of decline phase.

Francis Galton's (1869/1979) text, *Hereditary Genius*, is often cited as the first to describe expert development. Galton argued that natural talent was the main factor in developing expertise. From his analysis of eminent men of various professions and their families, he proposed that three factors were necessary to becoming an expert: innate ability, eagerness to work, and the power to accomplish a great deal of laborious work. Galton also claimed that there was an immutable limit on one's abilities, which no amount of education or exertion could overpass.

Galton's first factor, innate ability or natural talent, has gained much research attention but not a consensus of support. Research on child prodigies, as well as historical

accounts of experts' childhoods, has often been cited as proof that some individuals are born with a gift (Howard, 2008). Howe, Davidson, and Sloboda (1998) outline five criteria of natural talent in a domain as the following: talent is partly innate; full effects are not necessarily evident early on, but there are advance indications before later exceptional performance; early indications allow prediction of who is likely to excel; talent is domain specific; and talent is only exhibited in a minority of the population.

On the contrary, Ericsson and Charness (1994) reviewed the literature and found a lack of evidence for innate talent, with perhaps the exception of absolute pitch. In another study, the researchers concluded, with the exception of fixed genetic factors determining body size and height, there was limited support for innate constraints on attaining expertise (Ericsson, Roring, & Nandagopal, 2007). The notion of a giftedness or innate talent has also been questioned in the field of sociology. Based on longitudinal observation of swimmers, Chambliss (1989) concluded that success and expertise in sport are ordinary, simply a sum of all skills and contextual experiences. He emphasizes the qualitative differences that people do that are central to performance, rather what they innately possess.

While innate talent has been difficult to explain, focusing on factors like practice and training has led to a different explanation of acquisition of expertise. One early study shed light on the possibilities of practice (Ericsson, Chase, & Faloon, 1980). Although 7 ± 2 has been proposed as the magic number in terms of recall ability (Miller, 1956), these researchers trained a student to recall up to 80 digits after intense practice. This finding marks the importance of practice on performance, and it sparked further interest about how much practice one needs to be an expert.

When Simon and Chase (1973) reported that at least 10 years of practice was necessary to become an expert player, the connection between practice and expertise gained strength. Ericsson, Krampe, and Tesch-Römer (1993) extended these findings by comparing four groups of pianists, young and old amateurs and experts. Analysis of their practice routines demonstrated that the quantity of practice was not as important as the quality. More specifically, researchers suggested deliberate practice was crucial to attainment of expertise. This type of practice requires substantial effort; is not very enjoyable, but rather highly relevant; and will result in maximal improvement in performance (Ericsson et al., 1993). Deliberate practice involves informative feedback and provides many opportunities for repetition and correction of errors. These authors also put forward that a positive monotonic relationship exists between deliberate practice and performance.

Deliberate practice has been distinguished from actual practice (tournament performance or competition), as well as mere experience (playing for fun or leisure), in that it is a focused and effortful activity that one can engage in for only a limited time before fatigue and burnout set in (Ericsson, Roring, & Nandagopal, 2007). Although some transfer of practice to performance is expected from more leisurely domain-specific activities, experts want to maximize the efficiency of their training by engaging in challenging practice. As will be discussed later, there is some evidence that older experts show declines in practice and participation, without detrimental decline in performance. Some research suggests that older experts have learned to use their time wisely, thereby getting more out of less practice and performance time.

In Ericsson and Lehman's (1996) study on expert performance and adaptation to task constraints, performance was mediated by cognitive and perceptual-motor skills, as well as domain-specific physiological and anatomical adaptations. Ten or more years of daily deliberate practice was again stressed as necessary in attaining the highest level of human performance. While the 10-year rule of practice is largely accepted, some have pointed out that amount of practice time may vary depending on the nature of the sport, its history, and how many people play it around the world (Baker, Deakin, Horton, & Pearce, 2007). Additionally, deliberate practice researchers have often emphasized its solitary nature, which may be best for individual domains, but perhaps less so for those involving teams.

Besides innate talent and focused practice, other factors are relevant to acquiring expertise in a domain. Such factors include parental support, opportunity and access to resources, socioeconomic status, intellectual ability, and education. Domain-specific motivation (de Bruin, Rikers, & Schmidt, 2007) and passion (Vallerand et al., 2008) have also been linked to facilitating the extensive demands of engaging in deliberate practice. What continues to be evident - especially in the more recent literature - is that the development of expert performance is idiosyncratic and involves a complex interaction of factors related to each individual performer. Bilalić, McLeod, and Gobet (2007) proposed that with respect to varying performance, factors such as intelligence, practice, gender, experience, and age, are difficult to disentangle and should not be assessed separately. Similarly to how expertise is acquired, studies have also addressed when expert performance tends to peak.

Schulz and Curnow (1988) studied a number of sports that required speed and power (ex. running and swimming) as well as those more demanding in a cognitive-motor sense (ex. golf and baseball). What they found was that for the former, the peak age of expertise occurred in the early 20s, whereas sports that relied more heavily on cognitive-motor abilities peaked in late 20s and early 30s. These authors argued that because complex-motor skills take more time to acquire, expert sport domains where these skills are central tend to show a later peak in performance in terms of age. Domains such as music, art, and science would demonstrate peaks in one's 30s to 40s, similar to the domain of chess. Once performance peaks, of interest becomes how these levels of performance can be maintained by the aging expert and for how long.

Three primary explanations have been proposed for how experts maintain superior performance as they age, namely preserved differentiation, compensation (Salthouse, 1984), and selective maintenance (Krampe & Ericsson, 1996). The preserved differentiation or general factor explanation emphasizes innate talent even prior to skill development. This initial ability provides the expert with an edge throughout their entire career. While research supports that experts tend to excel in their domain as well as those closely related, no general advantage has been found across numerous domains (Ericsson & Charness, 1994).

The next perspective suggests that a compensatory mechanism allows for maintenance of superior performance while the expert experiences general age-related decline. Compensatory skills are domain-specific in order to offset declines that are inevitable with aging. This kind of logic is compatible with older experts possessing a

more efficient or refined strategy to their practice and performance, acquired over years of experience.

Lastly, the selective maintenance view emphasizes the importance of continued engagement in deliberate and domain-specific practice. Challenging practice, rather than mere involvement in the domain on a leisurely basis, is called for in order to maintain performance. Even deliberate practice in itself needs constant monitoring and revising in order to accommodate for optimal expert development over time. An expert would therefore not merely want to practice what they are skilled at, but rather would identify weaknesses in their performance and actively try to address them (Ericsson & Lehmann, 1996).

While rigorous practice seems like a promising avenue for prolonging one's expert career, decline does occur at some point. A number of researchers have addressed age-related decline in different expert domains, arriving at variable results. Studies of this kind all relate to the underlying question, "is aging kinder to the initially more able?"

1.3 Is Aging Kinder to the More Expert?

Research on aging experts has received a lot of attention in sport, but also in occupational settings such as air traffic control (Nunes & Kramer, 2009) and piloting (Morrow et al., 2009). What is generally of interest is whether the rate of decline as well as the profile over time is the same for experts versus those of more average ability. While some of the studies presented here focus more on intellectual performance, others are largely from the domain of athletics.

The assertion that “aging is kinder to the initially more able” suggests that those who excel at baseline will suffer less of an age-related decline over time. Regression to the mean is contrary to the kinder aging proposal. Discussed initially by Baltes, Nesselroade, Schaie, and Labouvie (1972) it suggests that both high and low ability groups will regress towards the overall average, such that the high ability group will suffer a decline, and the low ability group an increase, in ability over time. Their study found inconsistent age trends due to these statistical effects, and they stressed the importance of being aware of methodological issues. First, findings from diverse studies are summarized, followed by studies from the expert domain of athletics.

A number of researchers have addressed whether aging is kinder to the one who is initially more intelligent. Berkowitz and Green (1965) administered intelligence tests to American veterans and found that participants with above average IQ scores at baseline suffered the largest decline. They described an environment of inevitable, age-related “domiciliary” as being more restrictive for those who initially scored higher, compared to the lower ability group. Berkowitz and Green supported the disuse hypothesis of age-decline, which emphasizes lack of skill use rather than irreversible deterioration.

Riegel, Riegel, and Meyer (1967) measured intelligence, verbal abilities, attitudes, interests, and social conditions in their 55+ sample. From the results of a re-tested sample five years later, these researchers put forth that less able persons die earlier or are more likely to be seriously ill. Researchers warn against biased developmental trends concluded from samples that do not account for the attrition of non-survivors. They asserted that these individuals differ from those who are healthy enough to participate in the later testing.

In a later study, Riegel and Riegel (1972) found that age-related decline observed in a cross-sectional study was associated with increasing numbers of subjects exhibiting terminal drops. Decline with age was associated with a sudden drop in performance occurring five years before death, while long-term survivors showed little change in performance. Researchers speculated that participation in later studies is a powerful predictor of survival. The overall age difference in performance was an artifact caused by the continued participation at higher age levels of the surviving and cooperative subjects. Riegel and Riegel concluded that longitudinal studies were composed of the surviving and cooperative, whereas cross-sectional studies included subjects who are likely to “die within a few years and who are exhibiting terminal drops” (p. 9).

Christensen and Henderson (1991) studied eminent academics, blue-collar workers, graduate students, and students in trades to investigate whether intelligence and memory declined slower in those with higher ability. They found that pre-morbid intelligence and education contribute greater amounts of variance than chronological age. Yet, authors concluded that environmental stimulation, continued engagement in mental activities, and one’s physical health do not protect elderly individuals from decline in many areas of intellectual performance.

Gold and colleagues (1995) conducted a longitudinal study with 326 Canadian Army Veterans and found that there was relative stability in their scores over 40 years. They believed that intelligence during one’s young adulthood was the most important determinant of older adult performance. Brighter young men were also more likely to create a lifestyle that would help them retain verbal abilities as they aged. Overall, authors found their sample of community-based men in their mid-60s had experienced

relative stability in terms of intellectual performance. Studies that address whether aging is kinder to those initially more able have also been conducted in expert-only groups.

Krampe and Ericsson (1996) found that the amount of deliberate practice during later adulthood predicted the degree of maintenance of relevant pianistic skills for older experts. Older experts kept up a reduced, but still high, level of deliberate practice. Authors suggested that shifts in amount of deliberate practice with age could be in part due to increasing responsibility and less opportunity for practice in general. Further, they noted that older experts need to set aside more time for health and body care, if they want to maintain high levels of performance of their younger counterparts.

Ericsson (2000) also found that it is not so much age, as changes in practice with age that result in performance decline. Ericsson argued that reduced engagement in challenging activities and the decrease in intensity of maintained practice are key factors. Similarly, Masunaga and Horn (2001) concluded that if intensive, well-focused practice was taking place, age-related decline of the abilities did not necessarily occur in their sample of aging GO players.

Deary, Starr, and MacLennan (1998) proposed the differential aging hypothesis, namely that decline in ability scores over time will be lower in those with superior mental ability, more education, and higher social class at Time 1. Their findings showed that participants whose scores suffered a greater decrease, were significantly older, less educated, and had lower scores at baseline. The authors stressed that as exercise retains the body's fitness, the mind can go unfit if not challenged. The following studies address whether aging is kinder to those initially more expert in a number of athletic domains.

Horton, Baker, and Schorer (2008) noted two paradoxes with respect to expertise over the lifespan. First, that older experts are able to sustain a high level of performance in the face of overall decline in general capacities, and secondly, that they maintain it despite reduced involvement. Top athletes represent a cohort of individuals less influenced by disease and disuse, making them an ideal group for studies assessing age-related decline.

Stones and Kozma have been repeatedly acknowledged as the leaders in the field for identifying age-related changes in athletic performance (Baker, Horton, Pearce, & Deakin, 2005; Weir, Kerr, Hodges, McKay, & Starkes, 2002). Stones and Kozma (1984) studied short (anaerobic) and long (aerobic) running events, concluding that the aging process can be reflected in a linear or quadratic relationship, depending on the type of running event. Further, that cross-sectional and longitudinal studies addressing the age-related decline, paint a different picture of older experts. While longitudinal data reflect within-participant training effects, cross-sectional data do not. Inherent and cumulative training effects in longitudinal samples moderate the normal rate of age-related decline in performance (Young, Weir, Starkes, & Medic, 2008).

Further, Young and Starkes (2005) found that longitudinal data depicted the decline in running performance in a more linear fashion over time than cross-sectional data. Cross-sectional studies involve people who are intermittently active, which can lead to overestimated rates and extent of sports performance decline, versus longitudinal studies that are composed of individuals who train and compete across their lifetime.

Bortz and Bortz (1996) found a 0.5% yearly decline beginning at the age of 35 in their sample of professional runners, swimmers, and rowers. This estimate of decline was

also evidenced across different bodily systems such as VO₂ max (maximum oxygen consumption), DNA repair, and nail growth. Bortz and Bortz concluded that 0.5% could be a basic biomarker of the aging process.

In another study of young and old Master swimmers, practice and performance patterns were assessed (Weir, Kerr, Hodges, McKay, & Starkes, 2002). Practice patterns played a critical role in athletic performance and changed as swimmers aged to account for physiological changes. Older expert swimmers spent less time per week training and devoted more focus to endurance enhancement rather than strength during practice. Younger experts' practice patterns aimed more at improving endurance strength, speed, and power. Endurance focused training among older swimmers may help counter respiratory decline associated with aging.

Baker et al. (2005) examined the longitudinal performance decline in champion golfers. Golf performance peaks later because it is less constrained by the biological system and more dependent on acquired skills. Baker and colleagues found that golfers experience a constant decline in performance starting at the age of 35, but decline was slower than for the swimmers, runners, and rowers in the Bortzs' study. Age-related decline was 0.07% for the period between 35-49 and 0.26% between ages 51-60.

In a later study by the same research team, findings again showed that performance in golf can be maintained to a greater extent than in sport domains that rely more on biologically constrained skills (Baker, Deakin, Horton, & Pearce, 2007). Moreover, the rate and profile of decline in the face of advancing age was dependent on the skill under consideration. For example, greatest declines were evident in driving distance, a skill component that clearly needs bodily flexibility as well as adequate

strength. Authors suggest that golf may be similar to chess and piano when it comes to aging experts. A breakdown of trends for golfers at different stages in their PGA (Professional Golfers Association) careers was provided. Between ages 25 to 30, there was increased involvement in competitive play with highly improved performance, 30 to 43 was related to a consistent decline in number of rounds played but performance was maintained, while 43 to 50 was more characteristic of continued decline in number of competitive rounds with a corresponding gradual decline in performance. Explanations for competing less, but yet maintaining one's performance in the middle age range, suggested that golfers were still maintaining a high level of practice.

Young, Medic, Weir, and Starkes (2008) studied the performance of an elite middle-aged sample of runners in order to assess the contribution of age, as well as ongoing and past training factors. They found that while running performance slows with advancing age, specific training modifies this. Motivation, resources, and being free from injury are other factors that relate to training. Middle-aged running performance related most to training within the past five years. The authors concluded that support for merely early practice and late maintenance of performance is yet to be found.

In sum, while findings vary many studies seem to be in support of the aging process as kinder to those initially more able. Before moving into a review of chess research a brief summary of chess related facts will be provided.

1.4 FIDE

This section describes the World Chess Federation/ Fédération Internationale des Échecs (FIDE), as well as chess ratings. All information in this section is outlined on

FIDE's website (www.fide.com). FIDE was founded in 1924 in Paris and gained recognition by the International Olympics Committee as the International Sports Federation of Chess in 1999 (FIDE, 2011). FIDE is affiliated with chess federations in over 160 countries around the world. It acts as the supreme body responsible for chess organization as well as championships held worldwide.

Starting in the 1970s, FIDE started to issue lists of player statistics from international games. Today these rating lists are publicly accessible online and issued six times a year, with plans to be issued monthly. This increase is recent; when data were downloaded in July 2010 lists were issued 4 to 5 times per year.

Rating lists are available for a range of chess experts, both active and inactive players. A player is inactive when they have not participated in a FIDE-rated tournament in one year at the time of the issued list. Inactive players remain on the list for two more years. Players can be 'delisted' if they continue to remain inactive or if their ratings drop below the floor of 1200 points (FIDE, 2011).

FIDE links all players financially by collecting annual membership dues and mandating a standardized rating system. The system creates a "well-defined continuous hierarchy" among players, no matter where their geographic location (Puddephatt, 2008, p. 159). On this continuum, beginners rate in the 100s, while Grand Masters' ratings range from 2400 to 2800 points. A player is typically considered competitive if they are in the 1500 points range or above. Table 1 outlines rating ranges and their associated title categories in more detail.

Table 1

Rating System in Expert Chess: Rating Ranges and Title Categories

Rating range	Category
2600+	World Championship contenders
2400–2600	most Grandmasters (GM) and International Masters (IM)
2300–2400	FIDE Masters (FM)
2200–2300	FIDE Candidate Masters (CM), most national Masters
2000–2200	Candidate Masters, experts (USA)
1800–2000	Class A, category 1
1600–1800	Class B, category 2
1400–1600	Class C, category 3
1200–1400	Class D, category 4
below 1200	Novices

Note. Rating ranges and their associated categories as presented by FIDE. Fédération Internationale des Échecs (2011) in the FIDE Actual Handbook: Section C: General Rules and Recommendations for Tournaments. Retrieved from FIDE website: http://www.fide.com/handbook?option=com_handbook

Note that FIDE distinguishes between rating and title performance, the former being based on the player's result and the average rating of the opponent. Title performance depends on a more complex set of factors rendering comparison more difficult; as such, this study uses ratings. Some researchers point out that chess ratings are imperfect measures of underlying performance parameters (Glickman & Chabris, 1996). These authors suggest using a confidence interval to get a better estimate of a player's true rating, but this is uncommon in the literature. Further, ratings do not fully capture performance because they are uni-dimensional when chess performance involves a multidimensional skill set. Finally, Howard (2006) reminds us that rating archives contain many inaccuracies.

In order to build ratings, players must participate in tournament play. Chess tournaments are highly structured, involving globally accepted rules and associated penalties for violators. A player's rating determines whom they play in the tournament,

and this information is publicly displayed. The player's initial rating, their score (evaluation) from the game, and their opponents rating are all factored into their new rating. Thus, if a highly rated player loses to a player with a lower rating, the first player's rating will suffer more, than if the loss was to a player of more equal standing.

Players will often have a FIDE rating as well as a regional and/or national rating. As such, a FIDE rating does not reflect all games played, rather only those played at FIDE-recognized tournaments. One point of interest put forth by Jastrzemski, Charness, and Vasyukova (2006) is that ratings do not account for the level of participation of the player. Additionally, some suggest that older players tend to have decreased involvement in tournament play. Jastrzemski and colleagues argue that ratings for older players may not estimate chess skill as well as for younger players because of this drop in activity.

Roring and Charness (2007) also noted that "it is difficult to adjust for whether older or more highly skilled players might systematically play against different levels of competition, and it is possible that highly skilled players meet weaker opposition more often than vice versa" (p. 298). On the contrary, chess experts have described the ongoing pressures to study and compete if one has a high rating in their club's hierarchy (Puddephatt, 2008). In such a structured and monitored environment, it may be difficult to play the bare minimum of games.

1.5 Chess Research

The game of chess can be traced back to India, prior to 590 A.D. (Holding, 1985). Often referred to as the ultimate game of intellect, it is not surprising that chess has generated a lot of research in a number of fields. In psychology, researchers have used

chess to study intelligence (Grabner, Stern, & Neubauer, 2007), visuospatial abilities (Waters, Gobet, & Leyden, 2002), memory, perception (Charness, Reingold, Pomplun, & Stampe, 2001; Reingold, Charness, Schultetus, & Stampe, 2001), personality (Bilalić, McLeod, & Gobet, 2007; Vollstädt-Klein, Grimm, Kirsch, & Bilalić, 2010), the Flynn effect (Gobet, Campitelli, & Waters, 2002; Howard, 1999, 2005), expert performance (Roring & Charness, 2007), and even handedness (Gobet & Campitelli, 2007). More specifically, studies on expert chess have been conducted worldwide. Participants have spanned all age groups and skill levels, from prodigies to older experts (Horgan & Morgan, 1990; Howard, 2008; Frydman & Lynn, 1992).

There are a number of advantages that make chess suitable for such a variety of research. The domain is well defined and intellectually complex, which allows for the study of various cognitive processes. For the study of expertise, chess is useful because it quantifies performance via its standardized rating system. Longitudinal performance archives on performance, limited gatekeeper influence, and no need for participant recruitment are further advantages (Charness, 1992; Howard, 2006). Ratings are an objective measure of an entire population of top chess players, which negates concerns about sampling bias or need for inferential statistics. The next section presents studies from the chess expert literature in particular.

de Groot (1946/1978) is often cited as the first to study a group of chess experts. Moreover, this researcher compared two groups of experts, a less common design in the wealth of expert and novice studies. Researchers asked five Grand Masters and five Candidate Masters, presented with chess positions, to talk aloud while finding the best

move. Grand Masters used a similar number of moves and searched to similar depths, but they found better moves and more quickly.

In the 1970s, Simon and Chase presented two studies addressing perception and skill in chess (Chase & Simon, 1973; Simon & Chase, 1973). Simon and Chase estimated that a Master player possessed around 50,000 patterns of chess-related information that were stored and readily accessible. Researchers believed that thousands of hours of practice were responsible for remembering this many chunks (an analogy to how many words are in one's vocabulary after reading for a few decades). The perception study further alluded to familiarity, from years of practice, as playing a role in expertise. In fact, authors noted that approximately 75% of the moves are commonly written about in the chess literature, if one was extensively familiar with it. Speed of perceiving chunks was important in relation to chess skill, with the more expert players perceiving chunks faster.

The chunking hypothesis sparked further research, which included the Gobet and Simon study (1998). These researchers extended the chunking hypothesis to account for evidence of large retrieval structures in the long-term memory called templates. They believed chunks were organized around templates, which encode information acquired by skilled players over time. Gobet and Simon concluded that chess skill depends on recognition of familiar chunks of chess positions during play, as well as exploration and evaluation of possible moves and their consequences. They suggested expertise depended on the availability of memory for many frequently occurring patterns and the ability to conduct a highly selective search efficiently.

A later study conducted by Charness, Tuffiash, Krampe, Reingold, & Vasyukova (2005) found support for deliberate practice as the crucial ingredient to chess expertise. Two large diverse samples of chess experts retrospectively estimated the frequency and duration of their engagement in various chess-related activities over time. Researchers found that the number of hours a player dedicated to deliberate practice, followed by number of hours of tournament play, and years of private instruction best predicted chess skill. Further, that deliberate practice accounted for 40% of variance in skill level acquisition in expert chess players. These researchers concluded that independent study of books and positions is the most effective method of practice. In another study, Grabner, Stern, and Neubauer (2007) found that chess experience, current tournament activity, numerical intelligence, and personality factors accounted for 55% of the variability of playing strength in their sample.

While individual deliberate practice in chess is crucial to one's skill development, solitary practice is necessary but not sufficient. In their longitudinal study of 104 chess players of varied skill, Campitelli and Gobet (2008) found that recreational exposure to chess is also important. Further, that individual and group practice, plus factors unrelated to practice all, interact to achieve high level of performance in cognitive games such as chess. These authors suggest that social activities are more important than previously acknowledged (Charness et al., 2005), such that solitary learning should be coupled with rich environmental interaction. Besides development of chess expertise, some researchers have addressed the reasons for individual differences among top chess experts.

Charness and Gerchak (1996) questioned why Russian men dominate chess. They concluded that differences in performance are not due to cultural or innate reasons, but

rather due to participation rates. A country with a higher participation rate has more players who will get a chance to perform at high levels than one with a lower participation rate. Authors stressed the importance of accounting for group size before emphasizing group differences.

Some simply claim that women are worse chess players than men without providing much context on why this may be (Holding, 1985). Howard (1999, 2006) asserts that individual differences between male and female chess players are due to differences in innate abilities, intelligence, and visuospatial skills (1999, 2006). Howard concludes that men are simply more hard-wired for high chess achievement compared to women. Similar to the broader natural talent view, these findings are not supported by the bulk of the chess expert literature in general and some studies of visual memory in particular (Waters, Gobet, & Leyden 2002).

A recent study noted that 5% of registered tournament players and 1% of Grand Masters listed by FIDE were female (Maass, D'Ettolle, & Cadinu, 2008). These researchers have disputed Howard's conclusions about women in chess, based on how many more men participate at all levels. Authors highlight how women are clearly a visible minority, which can have an impact on performance in a domain. Bilalić and McLeod (2007) and Chabris and Glickman (2006) have also challenged Howard's conclusions. Bilalić and McLeod described how differences in participation rates, practice amounts, motivation, and interest between the two genders, better accounted for differences in chess performance, than intellectual or visuospatial abilities alone.

Chabris and Glickman (2006) investigated male dominance in chess by studying the statistics of 250,000 tournament players over a 13-year span. They concluded that

differences were due to more boys entering chess at all different skill levels in comparison to girls. Findings also showed that; men's rating were higher but no more variable than women's, boys and girls improved and dropped out at equal rates, boys began chess in greater numbers and at a higher level. These differences disappeared in countries where 50% of players were girls. In contrast, Howard (2009) found that male chess players still had higher ratings in a Georgian sample, although participation rates are nearly equal in this country.

Interesting findings came from a study of female and male chess players competing against each other online (Maass, D'Ettolle, & Cadinu, 2008). Female participants played several games against a male opponent. They were either told their opponent was male, falsely told they were female, or gender was not disclosed. When told their opponent was male, female players showed a drastic drop in performance and played more defensively. Female chess players performed nearly as well as their male counterparts, when gender was not disclosed. Finally, when led to believe they were facing a female opponent, female players performed equally well. Authors suggest the slight lag in performance, when gender was not disclosed, could be due to female players assuming they are facing a male opponent. This study demonstrates that minorities in expert chess, such as female players, can experience activation of such gender stereotypes. This can lead to internalized beliefs about oneself as inferior, which can certainly affect one's style of play and performance.

Cautious and defensive play has been associated with a prevention-orientation in which the player's goal is not to lose rather than to win. Aiming for at least a draw can also be seen as a preventative rather than promoting. Draws also introduce a potential for

strategic behavior that can secure a safer spot in a tournament as well as conserve a player's energy (Moul & Ney, 2009).

While a perspective involving motivation, participation, as well as marginalization may provide a more comprehensive explanation of the underperformance of women in chess, it draws some potential parallels to what older chess experts could experience later in their careers. While there is limited knowledge about ageism in chess or experiential accounts by older experts, a number of studies have addressed aging and chess expertise.

1.6 Chess Expertise and Aging

In 1960, Rubin reported on a survey about tournament performance and age of chess experts, finding that there was slight drop in skill around the age of 40. Rubin also suggested that older players drew more frequently. Draper (1963) analyzed World Championship chess matches and tournament performances over time. He found that expert chess players showed very little decline with age, until prior to death. Draper also found that these players won less, drew more, and lost about the same compared to earlier in their careers. Further, he speculated that older players were more content with draws to avoid risky moves. He also discussed the strains of an insecure life of an older chess expert and how trying to make a living off chess as one ages can erode performance due to limited study time.

Elo (1965) analyzed the tournament performance of 36 chess experts between 1885 and 1963 and produced a graph of their rating curves. Elo concluded that peak performance in this sample occurred around age 30 and older players regressed to

performance levels of their 20s. He proposed that after the age of 35, performance might decline because overall chess skill is a race between decreasing fluid intelligence after the age of 20 and increasing domain-specific knowledge through to experience. Later studies on chess expertise have found support for this peak age (Howard, 2005), while others suggest it occurs later in the early 40s (Roring & Charness, 2007).

Charness (1981) examined the relationship between age and chess skill with respect to problem solving in a group of chess players. He found that the critical aspects of chess players' performance do not relate to age, but more to skill. He further proposed that skill tends to trade-off with age in memory performance. Lastly, Charness suggested that since the participants all volunteered, perhaps this was a more active group and deficits due to age among chess players were underestimated.

In a 1999 study, Howard argued that humanity is becoming more intelligent, which was evidenced by the younger age demographic in expert chess since the 1970s. In a later study, Gobet, Campitelli, and Waters (2002) proposed a number of alternative explanations for this apparent trend. They suggested that younger experts, in comparison to their older counterparts, might have more access to the latest sport psychology literature and training technology. Moreover, that they are able to navigate through it more efficiently. They also highlight that expected highest achievement in a group is a function of its size and that perhaps more young people are playing (Charness & Gerchuk, 1996).

Recent research on chess expertise over time has benefitted from a database compiled by Howard (2006). This database consists of the statistics for over 60,000 top chess players dating back to the 1970s. Roring and Charness (2007) analyzed this

database using a multilevel modeling (MLM) approach in order to examine longitudinal change in chess skill of 5011 active players age 10-95. Although only 8.8% ($n = 417$) were over the age of 55, support for aging being slightly kinder to the initially more expert was concluded with acknowledgement of a milder decline past one's peak. While higher tournament activity predicted higher rating, there was an interaction with age for those more expert. The authors concluded that tournament activity had weaker effects on the ratings of older experts and that deliberate, solitary study is most important for this age group. Note that inactive players were excluded from analysis although statistics are available for this group. In the current study, inactive players will be included in order to get a more complete look at change over time.

Fair (2007) used best performance records to estimate age effects in swimming, track and field, and chess. He found that chess experts showed much less decline in performance than those in the other two domains. Fair suggested that estimated rates of decline could be benchmarks for other studies, because they were based on a large sample of chess players with a large age range, 35-94. Fair suggested that the 10-year rate of decline until the age of 80 is 4 %. His findings demonstrated that the function with the best rate was linear initially and quadratic as one aged.

Howard (2009) also utilized the database he constructed to assess development of chess expertise. He found that experts played more games and had greater increases in ratings than those who were less expert. He concluded that chess ratings vary positively with the number of games played, especially for expert players near the beginning of their careers. He also suggested that the strongest chess experts, namely the Grand Masters, showed precocity (entered domain younger and gained Grand Master title at

younger ages), faster acquisition of expertise (fewer years and games needed to achieve Grand Master title from when entered), and higher peak performance level.

Based on these studies, aging chess experts seem to avoid drastic declines in their performance ratings, but level of activity has been unaccounted for. Active participation in an expert domain is not without inactive periods, which serve their purpose in the development of skill and rating. The next section will discuss strategic participation in sports because this can affect performance standings. When a rating determines one's hierarchical identity in an elite community, like an international population of expert chess players, preservation of high ratings may be desirable.

1.7 Strategic Participation in Sport

Experts in various athletic domains consider when to train, how to train, and when to compete. Further, performers routinely consult coaches as well as other resources in order to maximize positive outcome, namely improving performance and winning. In many sports, including swimming, age-class categories determine who will be competing against whom. Fairbrother (2007) studied an American sample of 650 swimmers who ranked in the Top 10 quickest time records between 1993 and 2002. Top-ranked swimmers were lower than predicted by the regression equation after age 70, and the opposite was the case for the bottom-ranked swimmers. Swimmers' Top 10 rank was also higher when they transitioned into a new age category. From this came speculations that transition into an older age category provides the elite expert with an opportunity to maximize their performance.

A study conducted by Medic, Starkes, and Young (2007) provides some support for this last idea. These researchers proposed that expert athletes in track and field and swimming would utilize opportunities to compete at high levels in order to break records in the new age-class category they had just entered. Athletes in this sample participated more in national championships and broke more age-class records in the first two years following transition. Results suggested that elite experts, both young and old, cherish the chance for major accomplishment, are motivated to train for these opportunist competitions, and plan when to train and compete accordingly. Whether or not these windows of opportunity are available in a domain is an important parameter to consider because it may influence levels of practice and activity.

But what about sports with no age-class categories like chess? What kind of training and competition strategy would be most useful if age does not limit who plays whom? The chess federation is unique in that it does not organize tournaments by age. Instead, FIDE ratings determine competition, with the higher the rating, the more prestigious. Since age transitions do not create the same opportunities for chess experts, the goal is more likely to increase one's rating until it is time to preserve it. A rating preservation strategy could be viable for experts who are past their peak and still highly rated.

Interestingly, this may not be easy, as evidenced in one study where older chess experts described the pressures to compete in rated tournament games as inescapable and closely monitored (Puddephatt, 2008). Perhaps there exists a balance between rating preservation and obligation to the establishment for the aging expert? This qualitative study also demonstrated that some players internalized ratings as part of their core

identity. The extent that this occurred was associated with their level of skill, connection to outside interests, and position in the rating hierarchy. While all players are hierarchically organized in their respective clubs, top players' ratings are, in addition, constantly updated and followed online (de Bruin, Rikers, & Schmidt, 2007). This further makes ratings more central for expert players.

Chess players also face some unique challenges. Chess is very time-consuming and can strain the lives of players financially as well as interpersonally. The game of chess also has a low chance factor, which creates an atmosphere where accomplishment means more (Puddephatt, 2003). A player has no one else to blame for a loss but oneself. This can result in high sense of personal failure, especially if ratings are central to that player's identity. Some older experts may have a harder time adjusting to declines in their performance. On the other hand, if players are successful, they will be intrinsically motivated to continue to study and play. These qualitative findings provide a glimpse into the importance of ratings as well as pressure to perform in expert chess.

Although chess is often described as a cognitive domain, the stress, fatigue, and pressure that often accompany it can have a significant impact on physical and mental health (Holding, 1985). In a crude analysis of longevity among regional chess players, findings first shed light on the possibility that experts may experience more stress and shorter lives (Barry, 1969). In this early study, Barry suggested that a career outside of chess could act as a protective factor. This link to the role of outside interests, or the lack thereof, has been recently explored (Puddephatt, 2008). Perhaps older experts who do not define themselves solely in terms of their standings would have an easier time adjusting to their performance declines.

Motivation among chess experts, particularly with respect to strategic participation in chess has rarely been studied (de Bruin, Rikers, & Schmidt, 2007). One framework discussed by Higgins (1999) may be of some relevance. Higgins distinguishes between two motivational systems, one sensitive to gains (promotion) and the other to losses (prevention). While these systems described general life goals, they seem applicable to the motivation of an expert. A young expert may be particularly promotion-oriented; motivated to win games in order to develop their rating, while that same expert in later life may rather try to prevent losses and decline in their rating. Playing with the goal of minimizing losses in a competition can often lead to a more cautious or defensive approach (Moul & Ney, 2009). The notion that older experts draw more frequently has been raised in previous studies (Draper, 1963; Rubin, 1960), although the scientific rigor of these findings is questionable. Taking precaution to the extreme could range from participating on a minimal basis to not participating at all.

While inactivity may be strategic, it can also reflect different life priorities, stressors, less interest in competition in older age, illness or injury, or other potential factors. On the contrary, strategic inactivity may not be as simple to execute by an older expert in such a closely monitored community. In any case, to disregard player inactivity fails to provide a comprehensive view of age trends representative of the expert population as a whole.

1.8 Rationale

This study utilized multilevel modeling techniques to analyze a database of international chess experts, both active and inactive. The purpose was to examine the

trends in chess ratings in a recent five-year span, to see if aging is kinder to those initially more able. The main objective was to assess the mediating effects of various predictors (gender, entry age, elapsed age, baseline rating, activity status, and number of games played) and their interactions, as they related to overall ratings. Of particular interest was exploring the findings of Roring and Charness (2007) with respect to the relationship between tournament games and ratings for older experts. This was done by examining predictors like activity status and number of games played, as well as using a larger and more representative sample of players.

Finally, because chess has no age-group categories, high FIDE ratings are idealized, and players remain publicly listed even when inactive; what was also of interest, was whether inactive older experts suffered less decline in ratings. If this is the case, the overall implication is that ratings at a given age do not simply represent ability, but a combination of ability and the strategy to preserve higher ratings based on competitive activity. This study aimed to investigate these trends in an open-ended and exploratory manner.

Method

2.1 Participants

Data included statistics from 25 FIDE rating lists, encompassing a longitudinal span of five years (January 2006 to July 2010). Between 2006 and 2008, four lists were issued per year, namely January, April, July, and October. In 2009, five were available for the following months: January, April, July, September, and November. Lastly, 2010 lists include January, March, May, and July. These lists provide information on

demographics (gender, year of birth), ratings, and tournament activity for an international sample of chess players.

Cases omitted from the initial composite file include players with (1) no year of birth, (2) with ratings of 0, (3) inconsistency in year of birth within their player ID, (4) inconsistency in their name within player ID, (5) inaccurate or outlying year of birth, and (6) year of birth outside a range that included at least 100 different players for each year. Players with only one observation were included because this does not pose a problem in multilevel modeling analysis. Whatever data is available on a player contributes to estimating the best-informed individual growth curve as well as the average growth curve. The final sample consisted of 1,563,969 observations (cases) clustered within 99,608 players, 6 to 98 years of age.

2.2 Materials

Rating lists, in Notepad format, were downloaded from FIDE's website (<http://ratings.fide.com/download.phtml>) in July 2010. They were imported into SPSS 18.0 and incorporated into a larger composite file. The composite file was sorted according to player ID, with all observations for one player clustered together. This arrangement of sorting by cases is useful when working with repeated-measures data. From this format, steps were taken to eliminate duplicate cases and those not meeting inclusion criteria. The final database was analyzed using Mixed Linear Models in SPSS 18.0.

2.3 Variables

Player ID. Player ID was the subject variable. It was a unique number combination assigned by FIDE to each player.

Rating. Rating was the nested, dependent variable. The rating on a particular list is valid from the date of issue to the day prior to the next issued list (issued: Jan 2006, valid: Jan 1 - Feb 28/29, next issue: Mar 2006).

Chronological Age. Chronological age was an approximation of a player's age at the time of list publication. The extrapolated month and year of issue as well as the player's year of birth were used to arrive at this measure. This variable was the random component of this model. All of the following predictors were grand-mean centered. Centering transforms a variable's raw scores into deviations around a fixed point, such as the mean.

Entry Age. Entry age was a player's age at the time of when they first appeared on a rating list within the chosen period for analysis. This is not necessarily the age they entered FIDE lists (players rated prior to Jan 2006). This predictor and its interactions can provide information on trends across different age cohorts. The quadratic and cubic terms were computed for entry age in order to assess whether non-linear trends better account for growth in ratings. Since the development of ratings over a player's career is already a quadratic function, a cubic term is added to explore whether it can further explain the trends in data.

Elapsed Age. Elapsed age was a measure of how much time had elapsed from when a player first appeared on a rating list. This predictor can provide information on trends over time.

Gender. This variable identified a player as either male or female. Male players were assigned a value of 1, while female players were assigned a value of 2.

Activity Status. Activity status identified whether a player was active or inactive, with a value of 1 and 2, respectively. An active player played at least one FIDE-rated tournament game in the year prior to the issued list. An inactive player did not play any FIDE-rated tournament games for a minimum of one year, but remained on the list for two years after switching to this status.

Games Played. This variable was the number of FIDE-rated tournament games a player had participated in during the period prior to the issue date. Participating in FIDE-rated games is the only way a player can increase or decrease their current FIDE rating. Both, games played and activity status, provide information about a player's participation in FIDE-rated tournaments and expert chess.

Baseline Rating. Baseline rating was a player's rating on the first rating list they appeared on. This is not necessarily the rating they possessed when they first entered FIDE lists (players rated prior to Jan 2006).

2.4 Method of Analysis

Multilevel modeling (MLM), also known as hierarchical, mixed, or random coefficients modeling, has been increasingly used to assess developmental growth trajectories. More specifically, growth models can be used to assess the rate of change of a variable over time, with time as both a fixed and random effect. A number of recent texts describe the MLM approach and various applications (Bickel, 2007; Tabachnick & Fidell, 2007; Twisk, 2006).

MLM does not assume homogeneity of regression slopes. MLM does not require the assumption of independence of error (i.e. that residuals are uncorrelated). It can also handle missing observations and autoregressive tendencies in the data (Bickel, 2007). This allows for the inclusion of participants with incomplete data, which increases how representative the sample is. MLM also avoids the ecological fallacy, which makes inferences about individuals based on the aggregate statistics of the groups they belong to.

This method uses maximum likelihood estimation (MLE), in a series of iterations until the model converges. MLE tests the change in -2 log-likelihood (-2LL), which is an indicator of how much unexplained information there is after the model has been fitted. This value is used in accordance with the change in parameters and the chi-square distribution to make conclusions about the significance of the growing model. A Pseudo R^2 can also be computed to assess effect size. This measure is not an exact estimate of variance explained, such as R^2 , but rather a rough estimate. It will be reported in the form of a percentage, in order to capture the estimate more precisely.

Finally, MLM requires that all predictor variables be centered on their mean or some other significant value (taking each score and subtracting it from the mean of all scores for that variable). Centering allows for interpretation of the regression. Centering does not change the fit of the model (slope is not affected), but it does change the intercepts and associated statistics in terms of estimates of fixed effects.

2.5 Statistical Analysis

Initially, a null model was computed with Rating as the dependent variable, Player ID was the subject variable and centered Elapsed age was the repeated measure. The null model included only random variables, contained no additional predictor terms, and was the starting block for further comparisons of fit.

The repeated covariance structure used for all models was Scaled Identity (SI). Scaled Identity constraints residuals to have homogenous variance and be uncorrelated. Maximum likelihood (ML), rather than restricted maximum likelihood (REML), was used as the method of estimation because it allows for comparisons among different models. The covariance structure chosen for the random effects in the model was Unstructured (UN) because it imposes no constraints on the relationship between random components. While UN covers all bases, it also requires more parameters to estimate the model.

Following the null or random-intercept model, the next step involved setting the slope of the Chronological Age term as random. Predictors and their interactions were then added to the model one at a time. Each time the change in $-2 \log$ likelihood and parameters were calculated. In order to assess whether the change was significant, this calculation was compared to the critical values on the chi-square distribution. Different predictor variables were tested, but ultimately only those that contribute to the best fit remain.

The Results section describes the null model, a model without interactions (Model 10), a model with 2 and 3-way interactions (Model 18), and the final model with higher-order interactions (Model 20). For each, a discussion about the following is provided: -2

log likelihood value and χ^2 difference test, a Pseudo R^2 , estimates of fixed effects as well as the covariance parameters, or the random components of the model.

Results

3.1 Preliminary Findings

Players in the sample were 6 to 98 years of age ($M = 39.01$, $SD = 16.52$), at the time of data download. Year of birth spanned from 1912 to 2002 ($M = 1969.34$, $SD = 16.56$). Entry age ranged from 6 to 95 ($M = 37.20$, $SD = 16.43$).

Ratings ranged from 1212 to 2826, ($M = 2054.33$, $SD = 198.29$). Rating categories to keep in mind are as follows: elite experts or Grand Masters (2400+), experts or Masters (2200+), and those less expert (1200+). The majority of the sample fell into the least expert range ($n = 81,472$, 81.8%), then the expert group ($n = 15,781$, 15.8%), and finally the elite group ($n = 2,355$, 2.4%).

Games played on a list ranged from 0 to 110, ($M = 2.75$, $SD = 5.61$), with 0 values coming up most frequently.

3.2 Main Findings

The intraclass correlation coefficient (ICC) represents the proportion of the total variability of ratings that is attributable to player. A high ICC indicates that the majority of the variation occurs at the player level. The value of the ICC was 0.987. This value suggests that multilevel modeling is the appropriate because a high level of dependency exists among the repeated measures of one individual.

In multilevel modeling, terms and their interactions are added to the model in order to explore specific questions about the data at hand. Table 2 illustrates the chi-squared difference tests for each sequential step leading to an increasingly better fitting model.

Table 2

Comparison of Multilevel Models: Sequence of Model Parameters Entered

Model	Parameter added	-2 LL	df	χ^2 Difference Test ($M_n - M_{n+1}$)
M1 (Null)	Random intercept	15146102.3	3	
M2	Random intercept and slope	13856275.2	5	1289271.1***
M3	Gender	13856246.9	6	28.3***
M4	Activity status	13854115.3	7	2131.6***
M5	Games played	13849526.5	8	4588.8***
M6	Baseline rating	13816026.1	9	33500.4***
M7	Entry age	13813825.8	10	2200.3***
M8	Entry age ²	13808830.8	11	4995.0***
M9	Entry age ³	13807311.4	12	1519.4***
M10	Elapsed age	13807293.8	13	17.6***
M11	Entry age x Elapsed age	13796039.5	14	11254.0***
M12	Activity status x Entry age	13792804.7	15	3234.8***
M13	Activity status x Elapsed age	13791809.5	16	995.2***
M14	Games played x Entry age	13778740.7	17	13068.8***
M15	Games played x Elapsed age	13778544.4	18	196.3***
M16	Baseline rating x Entry age	13778391.2	19	153.2***
M17	Baseline rating x Elapsed age	13775738.4	20	2652.8***
M18	Baseline rating x Entry age x Elapsed age	13773872.4	21	1866.0***
M19	Games played x Baseline rating x Entry age x Elapsed age	13773783.2	22	89.2***

M20	Activity status x Baseline rating x Entry age x Elapsed Age	13773783.2	23	0.0 (NS)
-----	--	------------	----	----------

Note. This table provides a sequential outline of model parameters, -2 log likelihood values, and the chi-squared difference test to arrive at the best fitting model. M = Model. The following are the critical chi-square values for each probability level, $\chi^2 = 3.84, p < .05^*$, $\chi^2 = 6.63, p < .01^{**}$, and $\chi^2 = 10.83, p < .001^{***}$ for $df(1)$.

Tables 3 and 4 show the fixed and random terms for the null model or the random intercept model. This model tests whether there is a main difference between players in terms of mean ratings over time, which was significant.

Table 3

Estimates of Fixed Effects for the Null Model (DV = Rating)

Parameter	Estimate	SE	p	95% CI	
				LL	UL
Intercept	2006.01	.68	< .0001	2004.67	2007.35

Note. CI = confidence limit; LL = lower limit, UL = upper limit

Table 4

Estimates of Covariance Parameters for the Null Model (DV = Rating)

Parameter	Estimate	SE	p
Repeated Measures	603.42	.71	< .0001
Intercept	4.64E4	208.40	< .0001
[subject = ID]			

Note. CI = confidence limit; LL = lower limit, UL = upper limit

Next, a number of predictors were added to the random and fixed components of the model in a stepwise manner, in order to arrive at Model 10. Model 10, which contains no interaction terms, had 13 parameters and a -2 log likelihood value of 13807293.8. Using Model 9 for comparison, the $\chi^2 (df = 1) = 17.6, p < .0001$, a highly significant change. In order to assess the size of this effect, a Pseudo R^2 was computed and had a

value of 69.192 %. Tables 5 shows the fixed terms for Model 10.

Table 5

Estimates of Fixed Effects for Model 10 with no Interactions (DV = Rating)

Parameter	Estimate	SE	p	95% CI	
				-LL	UL
Intercept	2149.94	1.27	< .0001	2147.45	2152.44
Gender	-1.20	.35	.001	-1.89	-.52
Activity status	2.50	.05	< .0001	2.41	2.59
Games played	.18	2.69E-3	< .0001	1.77E-1	1.88E-1
Baseline rating	1.04	4.82E-3	< .0001	1.03	1.05
Entry age	-48.95	.87	< .0001	-50.65	-47.25
Entry age ²	1.12	.02	< .0001	1.07	1.16
Entry age ³	-6.87E-3	1.74E-4	< .0001	-7.22E-3	-6.53E-3
Elapsed age	.27	.06	< .0001	.14	.39

Note. CI = confidence limit; LL = lower limit, UL = upper limit

Table 5 illustrates a number of trends. Being a female player was associated with greater loss in ratings over time. Greater inactivity, in terms of activity status, was associated with less decrease in ratings over time. Games played, another measure of activity, was associated with less decrease in ratings over time. Higher baseline ratings were also associated with less decrement in ratings over time. Entry age has a curvilinear relationship with mean ratings as illustrated in Figure 1. Finally, greater elapsed age was associated with a less decrease in ratings over time. This means that the longer players are listed is associated with less decrease in ratings over time.

Figure 1. A Graph of Mean Ratings by Player's Entry Age

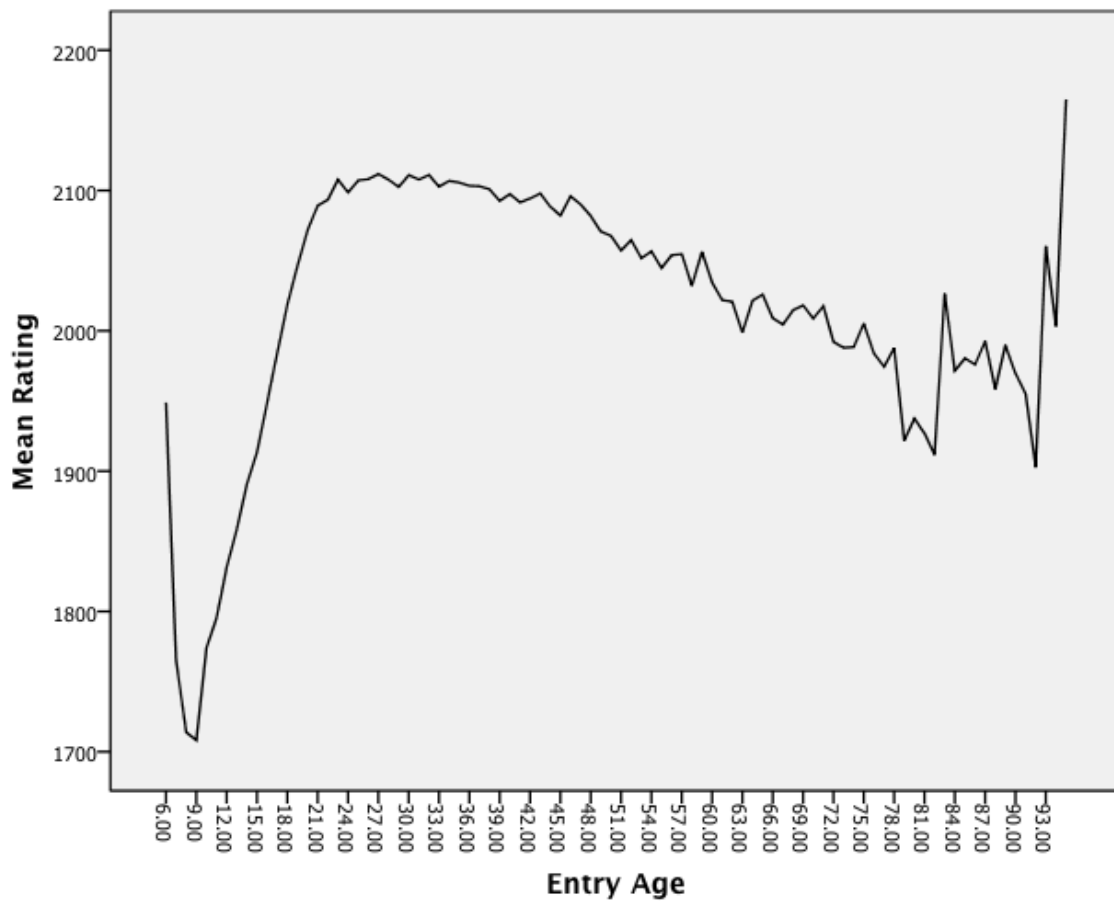


Figure 1. This graph demonstrates the curvilinear relationship between mean ratings and entry age. Mean ratings increase as player's skill develops in the early years of entry in expert chess. Mean ratings peak in the 30s, followed by slow decline. Beyond a linear and quadratic trend, a cubic trend can account for the rise in mean ratings in late years of entry.

Table 6 outlines the random portion of Model 10.

Table 6

Estimates of Covariance Parameters for Model 10 with no Interactions (DV = Rating)

Parameter	Estimate	SE	<i>p</i>
Repeated Measures Intercept + Chronological age [subject = ID]	185.90	.23	< .0001
UN (1,1)	3.09E5	1702.96	< .0001
UN (2,1)	-9151.56	50.32	< .0001
UN (2,2)	359.21	1.82	< .0001

Note. CI = confidence limit; *LL* = lower limit, *UL* = upper limit

As evident in Table 6, players varied significantly in terms of their intercepts, $\text{var}(u_{0j}) = 3.09E5$, $p < .0001$. Time trajectories or slopes also significantly differed across players, $\text{var}(u_{1j}) = 359.21$, $p < .0001$. The $\text{cov}(u_{0j}, u_{1j}) = -9151.56$, $p < .0001$ indicates that as intercepts increase, slopes decrease. As such, players with higher ratings on average, demonstrate less drastic changes in their ratings over time.

Model 18 included 2-way and 3-way interactions, had 21 parameters, and a -2 log likelihood value of 13773872.4. Using Model 17 for comparison, the $\chi^2(df = 1) = 1866.0$, $p < .0001$, a highly significant improvement. In order to assess the size of the effect of the addition of interaction terms in Model 18, a Pseudo R^2 was computed and had a value of 69.414%. Tables 7 and 8 show the fixed and random terms for Model 18.

Table 7

Estimates of Fixed Effects for Model 18 with Interactions (DV = Rating)

Parameter	Estimate	SE	<i>p</i>	95% CI	
				<i>LL</i>	<i>UL</i>
Intercept	2051.79	1.46	< .0001	2048.93	2054.65
Gender	-2.06	.35	< .0001	-2.74	-1.38
Activity status	1.84	.05	< .0001	1.75	1.94
Games played	.13	2.79E-3	< .0001	.12	.13

Baseline rating	.96	6.55E-3	< .0001	.95	.98
Entry age	-24.26	.92	< .0001	-26.06	-22.45
Entry age ²	.67	.02	< .0001	.63	.72
Entry age ³	-6.06E-3	1.78E-4	< .0001	-.01	-.01
Elapsed age	-.11	5.91E-2	.071 (NS)	-.22	.01
Entry age*Elapsed age	-.36	.36E-3	< .0001	-.37	-.35
Activity status*Entry age	.15	2.99E-3	< .0001	.14	.16
Activity status*Elapsed age	1.34	3.68E-2	< .0001	1.27	1.42
Games played*Entry age	-1.87-E2	1.60E-4	< .0001	-.019	-.018
Games played*Elapsed age	.03	2.03E-3	< .0001	2.21E-2	3.01E-2
Baseline rating*Entry age	-8.00E-4	4.13E-4	.053 (NS)	-1.61E-3	9.27E-6
Baseline rating*Elapsed age	-1.44E-2	2.82E-4	< .0001	-1.49E-2	-1.38E-2
Baseline rating*Entry age*Elapsed age	5.15E-4	1.19E-5	< .0001	4.91E-4	5.38E-4

Note. CI = confidence limit; LL = lower limit, UL = upper limit

Many of the significant fixed effects in Model 10 also held true for Model 18. One difference in Model 18 was the entry age predictor, which had decreased quite a bit, thus contributing less to the explanation of within-in group differences. Elapsed age in Model 18 had also undergone change, as now it indicated that the longer the player appears on the list, the more their rating suffered over time.

The significant interactions in Model 18 are described next. Elapsed age interacted with entry age, such that older players, suffered greater loss in ratings over time. Activity status interacted with entry age, such that more inactive older cohorts

showed less decrease in ratings over time. A similar trend was found for the activity status by elapsed age interaction. Games played interacted negatively with entry age, such that the older players that played more games, suffered greater drops in rating. In contrast, games played interacted with elapsed age, such that the more games played and the longer one was listed, the less the rating suffered. Baseline ratings interacted with elapsed age, but not entry age. The higher the baseline rating and the longer one was listed, was associated with greater loss in ratings over time. Finally, the 3-way interaction between baseline rating, entry age, and elapsed age was significant. The higher the baseline rating and the older the player and the longer listed; the lower the loss in ratings over time. Table 8 outlines the statistics for the random portion of Model 18.

Table 8

Estimates of Covariance Parameters for Model 18 with Interactions (DV = Rating)

Parameter	Estimate	SE	<i>p</i>
Repeated Measures	184.56	.23	< .0001
Intercept + Chronological age [subject = ID]			
UN (1,1)	2.71E5	1479.83	< .0001
UN (2,1)	-7718.35	42.75	< .0001
UN (2,2)	297.22	1.54	< .0001

Note. CI = confidence limit; *LL* = lower limit, *UL* = upper limit

Similar to the results for the random component of Model 10, Table 8 illustrates the same relationships in Model 18. Players varied significantly in terms of their intercepts, $\text{var}(u_{0j}) = 2.71E5$, $p < .0001$. Time trajectories or slopes also significantly differed across players, $\text{var}(u_{1j}) = 297.22$, $p < .0001$. Finally, the $\text{cov}(u_{0j}, u_{1j}) = -7718.35$, $p < .0001$ again indicates that as intercepts increase, slopes decrease.

Model 20 included higher-order interactions, had 23 parameters, and a -2 log likelihood value of 13773783.2. Using Model 18 for comparison, the $\chi^2(df = 2) = 89.2$,

$p < .0001$. In order to assess the size of the effect of the addition of the interaction terms in Model 20, a Pseudo R^2 was computed and had a value of 69.416%. Table 9 outlines the fixed terms of Model 20.

Table 9

Estimates of Fixed Effects for the Final Model with Higher-Order Interactions (DV = Rating)

Parameter	Estimate	SE	p	95% CI	
				LL	UL
Intercept	2051.84	1.46	< .0001	2048.99	2054.70
Gender	-2.13	.35	< .0001	-2.81	-1.45
Activity status	1.84	.05	< .0001	1.75	1.94
Games played	.13	2.79E-3	< .0001	1.21E-1	1.32E-1
Baseline rating	.96	.01	< .0001	.95	.97
Entry age	-24.27	.92	< .0001	-26.08	-22.47
Entry age ²	.67	.02	< .0001	.63	.72
Entry age ³	-.01	1.78E-4	< .0001	-6.41E-3	-5.71E-3
Elapsed age	-.11	.06	.070 (NS)	-.22	.01
Entry age*Elapsed age	-.36	3.61E-3	< .0001	-.37	-.35
Activity status*Entry age	.15	2.99E-3	< .0001	1.47E-1	1.59E-1
Activity status*Elapsed age	1.35	.04	< .0001	1.27	1.42
Games played*Entry age	-1.90E-2	1.65E-4	< .0001	-1.94E-2	-1.87E-2
Games played*Elapsed age	.03	2.03E-3	< .0001	.02	.03
Baseline rating*Entry age	-8.53E-4	4.13E-4	.039	-1.66E-3	-4.39E-5
Baseline rating*Elapsed age	-.01	2.83E-4	< .0001	-1.48E-2	-1.37E-2
Baseline rating*Entry age	5.02E-4	1.20E-5	< .0001	4.79E-4	5.26E-4

age*Elapsed age Games	4.73E-6	5.05E-7	< .0001	3.74E-6	5.72E-6
played*Baseline rating*Entry age*Elapsed age Activity	-1.26E-6	1.28E-5	.922 (NS)	-2.63E-5	2.38E-5
status*Baseline rating*Entry age*Elapsed Age					

Note. CI = confidence limit; *LL* = lower limit, *UL* = upper limit

In Model 20, many of the significant predictors from Model 18 remained unchanged. An exception was the predictor of elapsed age, which in Model 20 was no longer significant. This type of switch to non-significance often occurs as additional predictors are added into the model. All interactions from Model 18 were also significant in Model 20. In addition, the interaction between baseline rating and entry age was now significant, such that the older the player and the higher their baseline rating, the more ratings suffered over time. Lastly, two additional 4-way interactions were tested in Model 20, one of which was significant. Those who played more games, had higher baseline ratings, were older at time of entry, and appeared on rating lists for a longer time; showed a lower decrease in ratings over time. Table 10 outlines the random portion of Model 20.

Table 10

Estimates of Covariance Parameters for the Final Model with Higher-Order Interactions (DV = Rating)

Parameter	Estimate	<i>SE</i>	<i>p</i>
Repeated Measures	184.55	.23	< .0001
Intercept + Chronological age [subject = ID]			
UN (1,1)	2.71E5	1480.67	< .0001
UN (2,1)	-7716.51	42.75	< .0001
UN (2,2)	297.04	1.54	< .0001

Note. CI = confidence limit; *LL* = lower limit, *UL* = upper limit

The results in Table 10 depict the same pattern of relationships between individual players and time, as do Tables 6 and 8. Players varied significantly in terms of their intercepts, $\text{var}(u_{0j}) = 2.71\text{E}5, p < .0001$. Time trajectories or slopes also significantly differed across players, $\text{var}(u_{1j}) = 297.04, p < .0001$. Finally, the $\text{cov}(u_{0j}, u_{1j}) = -7716.51, p < .0001$ again indicates that as intercepts increase, slopes decrease. This means players with higher initial ratings experience less drastic changes in their ratings over time. Finally, Figure 2 demonstrates mean change from baseline rating by elapsed age for three different age groups and three rating categories.

Figure 2. A Collapsed Graph of Mean Change from Baseline Rating by Elapsed Age for Three Age Groups and Rating Categories

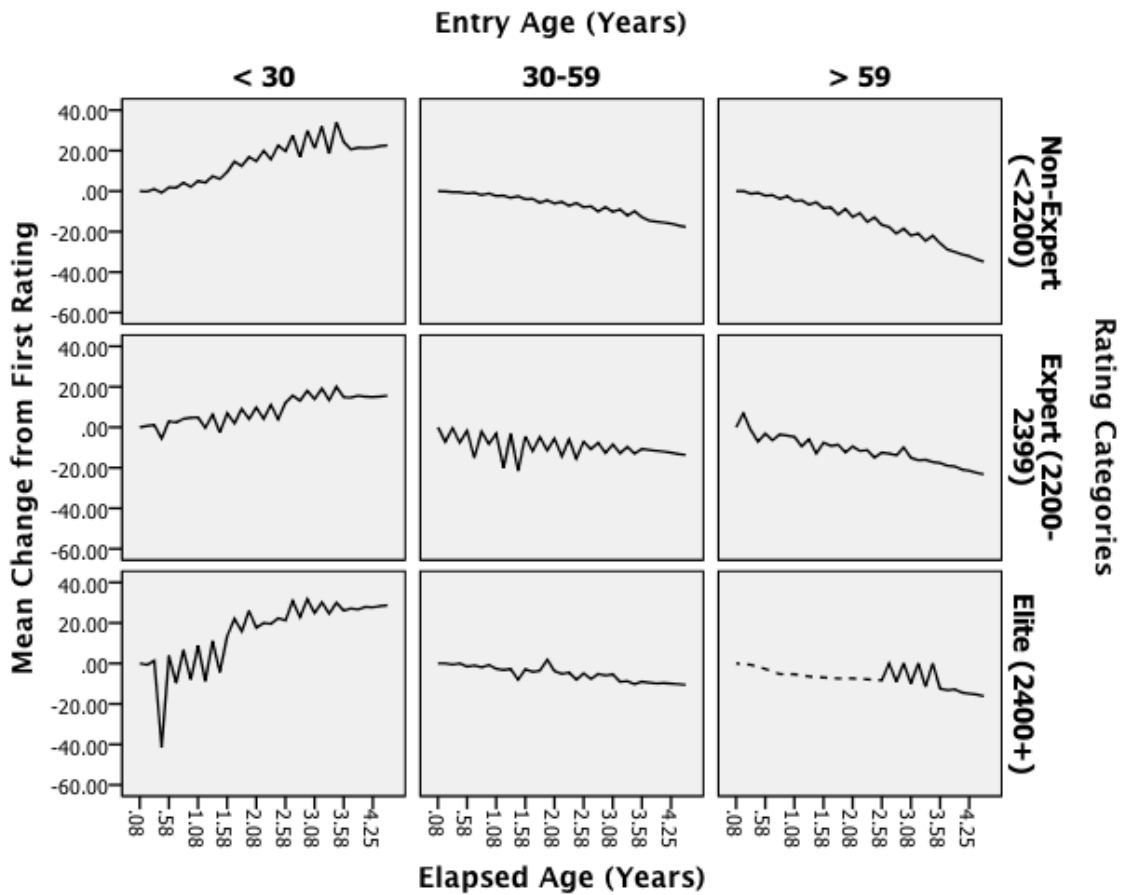


Figure 2. This graph depicts the mean change from baseline rating by elapsed age for different age groups as well as rating categories. Age groups for entry age were less than 30 years of age, 30 to 59, and 60 years of age and older. The three rating categories included the following: non-expert or less expert (<2200 points), experts (2200-2399) and elite experts (2400+ points). This graph illustrates that the youngest age group experiences highest mean change in ratings, no matter what the rating category. The middle age group experiences the least mean change in ratings, with the expert group showing most variability. Finally, the older age group shows a steep decline in the least

expert group, and a milder decline for those in the expert group. The elite experts in the older age group continue to demonstrate similar mean changes to the elite, middle-aged cohort. It is evident that the ratings of the oldest, elite group suffer the least, in terms of decline in ratings. Age categories were used in this graph in order to illustrate the 3-way interaction. Age categories were not utilized in the MLM analysis.

Discussion

4.1 Summary of the Findings

Multilevel modeling was used to analyze this dataset because it is appropriate when there is established theory, as well as a small amount of unrelated predictors. While this study was exploratory, the literature in this area supports aging being kinder to those more expert. This study included a small number of predictors and an extensive amount of players and observations. All predictors were grand-mean centered, since main effects and interactions can often be correlated. Because measures were taken to adhere to how MLM is to be used, we can be sure that the findings reflect the trends in this data.

A wide age range was assessed (6 to 98). This sample allowed for the assessment of age cohorts as well as age change over time. Ratings ranged from close to the minimum allowed by FIDE (1212 points) up to the most skilled of players (2826 points). Most players were in the less expert category (81%), while fewer made up the expert (15.8%) and elite expert (2.4%) categories. Games played ranged between 0 and 110, with an average of 2.75 games. This demonstrates that inactivity is very much a part of expert chess and needs to be accounted for when working with chess performance records.

Based on the ICC value, differences between players accounted for 98.7% of the variance in ratings. Of the 1.3% remaining within-in player variance, the Pseudo R^2 estimate indicates the model explains 69.42% of the within-in player difference in ratings. Being of female gender was significantly associated with greater loss in ratings. This is not surprising given the previously discussed studies addressing marginalized status of women in chess and lower participation rates. Greater inactivity, more games played, and higher baseline ratings, were all associated with less loss in ratings over time. What is of interest is that both inactivity and activity (more games played) were associated with less loss in ratings for players overall. This may reflect the reality that different pathways, both active and inactive, contribute to slower rating decline.

The higher-order interactions addressed the main objectives of this study. The 3-way interaction showed that, the higher the baseline rating, the older the age of entry, and the longer the player was listed, the less loss in mean ratings over time. These conclusions are in line with those of Roring and Charness (2007) who also found that aging was slightly kinder to those more expert. The findings also echo the conclusions of other researchers addressing chess and aging, particularly with respect to mild performance decline (Charness, 1981; Draper, 1963; Fair, 2007; Howard, 2009; Rubin, 1960).

Unlike the Roring and Charness study though, a significant positive effect of games played, on ratings over time, was found for more expert players. This has important implications for aging chess experts. In their study, Roring and Charness (2007) concluded playing more games had a positive effect on ratings, but this effect was small for older experts. Based on this finding, authors suggested older experts focus

primarily on solitary practice. The present findings, on the contrary, demonstrate that older experts still benefit from participating in tournament play. In particular, older players who are initially more able (higher baseline ratings), play more games, and who appear on rating lists for a longer time period, experienced less loss in terms of mean ratings over time. This study was unique in that it analyzed a near population of expert players in a recent time period. Also, this sample was more representative because it included inactive players and players with a single entry.

Although the objectives were exploratory, there was an alternative argument presented throughout the introductory sections, about how older experts could use the inactivity status option to their advantage. The findings did not show support for this line of reasoning and instead demonstrated that tournament involvement has positive effects on ratings over time.

The positive effect of games played for the ratings of older elites is not surprising. These players are unique because of their ongoing commitment to the study, practice, and play (Roring & Charness, 2007). This commitment has likely developed because of the ongoing success they have experienced and the vast repertoire of small skills they have developed over time (Chambliss, 1989). This top stratum is also most likely to gain the greatest intrinsic and social rewards from continued participation. Expert players also demonstrate high levels of engrossment, described as a deep focus and complete surrender to the flow of the game. Engrossment in an activity has been associated with creative, efficient, and skillful play, resulting in high performance and long-term dedication to the sport. Finally, qualitative findings have shed light on the pressures to stay active, as reported by aging, highly rated club players (Puddephatt, 2008).

4.2 Limitations of the Study

There are a few limitations, which have also been identified and acknowledged in past research. The rating list records are real-world data and could contain inaccuracies (Howard, 2006), although measures were taken to remove problematic cases. Ratings are uni-dimensional while performance and skill are multi-faceted. Ratings also do not reflect participation rates (Glickman & Chabris, 1996).

Concluding that aging is kinder to those initially more able in chess could be a reflection of the following factors. It has been suggested that there is selective attrition in expert chess over time and that only the top remain (Roring & Charness, 2007). Others have suggested that the most expert players may be selective about their competition and try to meet weaker opponents (Jastrzemski, Charness, & Vasyukova, 2006). The way FIDE calculates ratings also complicates matters. Players at the top of the rating hierarchy improve and decline at slower rates, due to a more conservative weighing factor in the FIDE calculation. All of these issues relate to older experts ratings being protected at high levels and not truly reflecting the declines that may be occurring.

A future follow-up study could assess these trends on the national level by utilizing the records kept by the Canadian Chess Federation (CCF). The CCF keeps ongoing records of performance ratings and player ratings. Player ratings are similar to the FIDE ratings used in this study, such they arrange players in a hierarchical order based on an overall value. Performance ratings, on the other hand, refer to the rating level displayed in the actual tournament. Performance ratings may be more accurate indicators of actual ability, since they are more variable and reflect performance level in the present.

An aging expert for example, may still have a high overall rating but may show a significant decline in performance ratings during tournaments. While performance ratings may continue to suffer, there may be a time lag before this is reflected in their overall rating. As such, exploring CCF's performance ratings and how they are predicted by the independent variables assessed here, may provide novel insights, particularly for aging experts.

4.3 Implications

Chess continues to be a domain of expertise in which expert players experience a milder decline in terms of their ratings. Ratings as a true reflection of performance, as well as the FIDE rating system as a whole, have various limits as previously outlined. These caveats in chess and aging research have been acknowledged, but ratings continue to be used as the best available measure of performance in this domain. As such, they have been utilized in this study, but with a more recent and representative sample. The predictors in this model were also able to explain a large portion of the within-in group variance in ratings.

The findings demonstrate that playing more games is associated with less decrement in performance ratings for those initially more able. The critical implication here is that older experts still benefit from tournament play and can positively affect the development of their ratings by playing – or minimize the drop in ratings. While solitary practice has been emphasized and certainly plays a key role, actual tournament participation is valuable!

4.4 Conclusions

Clearly, a complex set of factors explain aging and changes in performance ratings among expert chess players. The present findings demonstrate that, for those initially more able in expert chess, the level of involvement in tournament play does have an effect. In fact, the more games played and the longer one is involved in tournament play, the higher the ratings for those initially more expert. While strategic inactivity is still an alternative players may choose to maintain high ratings, these trends demonstrate the benefits of continued activity. In conclusion, as exercise keeps the body fit, playing more games keeps performance ratings high - especially for those initially more able.

References

- Baker, J., Deakin, J., Horton, S., & Pearce, G. W. (2007). Maintenance of skilled performance with age: A descriptive examination of professional golfers. *Journal of Aging and Physical Activity, 15*(3), 300-317.
- Baker, J., Horton, S., Pearce, G. W., & Deakin, J. (2005). A longitudinal exam of performance decline in championship golfers. *High Ability Studies, 16*, 179–185.
doi:10.1080/13598130600617928
- Baltes, P. B., Nesselroade, J. R., Schaie, K. W., & Labouvie, E. W. (1972). On the dilemma of regression effects in examining ability-level-related differentials in ontogenetic patterns of intelligence. *Developmental Psychology, 6*(1), 78-84.
doi:10.1037/h0032329
- Barry, H. (1969). Longevity of outstanding chess players. *Journal of Genetic Psychology, 115*(1), 143-148.
- Berkowitz, B., & Green, R. F. (1965). Changes in intellect with age: V. differential changes as functions of time interval and original score. *Journal of Genetic Psychology, 107*(2), 179-192.
- Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression*. New York, NY: The Guilford Press.
- Bilalić, M., & McLeod, P. (2007). Participation rates and the difference in performance of women and men in chess. *Journal of Biosocial Science, 39*(5), 789-793.
doi:10.1017/S0021932007001861

- Bilalić, M., McLeod, P., & Gobet, F. (2007). Does chess need intelligence? A study with young chess players. *Intelligence*, *35*, 457-470. doi:10.1016/j.intell.2006.09.005
- Bilalić, M., McLeod, P., & Gobet, F. (2008). Expert and “novice” problem solving strategies in chess: Sixty years of citing de Groot (1946). *Thinking & Reasoning*, *14*(4), 395-408. doi:10.1080/13546780802265547
- Bortz, W. M., & Bortz, W. M. (1996). How fast do we age? Exercise performance over time as a biomarker. *Journal of Gerontology*, *51*, 223-225. doi:10.1093/gerona/51A.5.M223
- Campitelli, G., & Gobet, F. (2008). The role of practice in chess: A longitudinal study. *Learning and Individual Differences*, *18*(4), 446-458. doi:10.1016/j.lindif.2007.11.006
- Chabris, C. F., & Glickman, M. E. (2006). Sex differences in intellectual performance: Analysis of a large cohort of competitive chess players. *Psychological Science*, *17*(12), 1040-1046. doi:10.1111/j.1467-9280.2006.01828.x
- Chambliss, D. F. (1989). An ethnographic report on stratification and Olympic swimmers. Retrieved from The Mundanity of Excellence website: www.scidb.com/doc/2926754/The_Mundanity_of_Excellence
- Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology; Human Perception Performance*, *7*, 467-476. doi:10.1037/0096-1523.7.2.467

- Charness, N. (1992). The impact of chess research on cognitive science. *Psychological Research/Psychologische Forschung*, 54(1), 4-9. doi:10.1007/BF01359217
- Charness, N., & Gerchak, Y. (1996). Participation rates and maximal performance: A log-linear explanation for group differences, such as Russian and male dominance in chess. *Psychological Science*, 7(1), 46-51. doi:10.1111/j.1467-9280.1996.tb00665.x
- Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition*, 29(8), 1146-1152.
- Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology*, 19(2), 151-165. doi:10.1002/acp.1106
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4(1), 55-81. doi:10.1016/0010-0285(73)90004-2
- Chi, M. T. H. (2006) Two approaches to the study of experts' characteristics. In: K.A. Ericsson, N. Charness, P.J. Feltovich and R.R. Hoffman, Editors, The Cambridge handbook of expertise and expert performance, Cambridge University Press, Cambridge, U.K. (2006), pp. 21–30.
- Christensen, H., & Henderson, A. S. (1991). Is age kinder to the initially more able? A study of eminent scientists and academics. *Psychological Medicine: A Journal of*

Research in Psychiatry and the Allied Sciences, 21(4), 935-946.

doi:10.1017/S0033291700029925

Deary, I. J., Starr, J. M., & MacLennan, W. J. (1998). Is age kinder to the initially more able? Differential ageing of a verbal ability in the Healthy Old People in Edinburg study. *Intelligence*, 26(4), 357-375. doi:10.1016/S0160-2896 (99) 00005-7

de Bruin, A. B. H., Rikers, R. M. J. P., & Schmidt, H. G. (2007). The influence of achievement motivation and chess-specific motivation on deliberate practice. *Journal of Sport & Exercise Psychology*, 29(5), 561-583.

de Groot, A. D. (1946/1978). *Thought and choice in chess*. (2nd ed). Oxford, England: Mouton.

Draper, N. R. (1963). Does age affect master chess? *Journal of the Royal Statistical Society*, 126, 120-127.

Elo, A. E. (1965). Age changes in master chess performances. *Journal of Gerontology*, 20, 289-299.

Ericsson, K. A. (2000). Expertise in interpreting: An expert-performance perspective. *Interpreting*, 5(2), 189-233. doi:10.1075/intp.5.2.08eri

Ericsson, K. A., & Charness, N. (1994). Expert performance; Its structure and acquisition. *American Psychologist*, 49(8), 725-747. doi:10.1037/0003-066X.49.8.725

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). *Acquisition of memory skill*. *Science*, *208*, 1181-1182. doi:10.1126/science.7375930

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273-305. doi:10.1146/annurev.psych.47.1.273

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363-406. doi:10.1037/0033-295X.100.3.363

Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Studies*, *18*, 3-56. doi:10.1080/13598130701350593

Fair, R. C. (2007). Estimated age effects in athletic events and chess. *Experimental Aging Research*, *33*(1), 37-57. doi:10.1080/03610730601006305

Fairbrother, J. T. (2007). Prediction of 1500-m freestyle swimming times for older masters All-American swimmers. *Experimental Aging Research*, *33*(4), 461-471. doi:10.1080/03610730701525402

Fédération Internationale des Échecs. (2011). FIDE Actual Handbook. Retrieved from FIDE website: http://www.fide.com/handbook?option=com_handbook

Fédération Internationale des Échecs. (2010). *FIDE Rating List* [25 data files]. Retrieved from FIDE website: <http://ratings.fide.com/download.phtml>

- Frydman, M. & Lynn, R. (1992). The general intelligence and spatial abilities of gifted young Belgian chess players. *British Journal of Psychology*, *83*, 233-235.
- Galton, F. R. S. (1869/1979). Hereditary genius: An inquiry into its laws and consequences. D. Appleton and Company, New York.
- Glickman, M. E., & Chabris, C. F. (1996). Using chess ratings as data in psychological research. In Department of Mathematics and Statistics, Boston University 1996 Boston, MA: Department of Mathematics and Statistics, Boston University Unpublished manuscript, available from www.wjh.harvard.edu/cfc/Glickman1996.pdf
- Gobet, F., & Campitelli, G. (2007). The role of domain-specific practice, handedness, and starting age in chess. *Developmental Psychology*, *43*(1), 159-172. doi:10.1037/0012-1649.43.1.159
- Gobet, F., Campitelli, G., & Waters, A. J. (2002). Rise of human intelligence: Comments on Howard (1999). *Intelligence*, *30*, 303-311. doi:10.1016/S0160-2896(02)00083-1
- Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory* *6*(3), 225-255. doi:10.1080/741942359
- Gold, D. P., Andres, D., Etezadi, J., Arbuckle, T., Schwartzman, A., & Chaikelson, J. (1995). Structural equation model of intellectual change and continuity and predictors of intelligence in older men. *Psychology and Aging*, *10*(2), 294-303. doi:10.1037/0882-7974.10.2.294

- Grabner, R. H., Stern, E., & Neubauer, A. C. (2007). Individual differences in chess expertise: A psychometric investigation. *Acta Psychologica, 124*(3), 398-420. doi:10.1016/j.actpsy.2006.07.008
- Higgins, E. T. (1999). Promotion and prevention as a motivational duality: Implications for evaluative processes. In S. Chaiken, & Y. Trope (Eds.), (pp. 503-525). New York, NY, US: Guilford Press.
- Holding, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Horgan, D. D., & Morgan, D. (1990) Chess expertise in children. *Applied Cognitive Psychology, 4*, 109-128. doi:10.1002/acp.2350040204
- Horton, S., Baker, J., & Schorer, J. (2008). Expertise and aging: maintaining skills through the lifespan. *European Review of Aging and Physical Activity, 5*, 89-96. doi:10.1007/s11556-008-0034-5
- Howard, R. W. (1999). Preliminary real-world evidence that average human intelligence really is rising. *Intelligence, 27*(3), 235-250. doi: 10.1016/S0160-2896(99)00018-5
- Howard, R. W. (2005). Objective evidence of rising population ability: A detailed examination of longitudinal chess data. *Personality and Individual Differences, 38*(2), 347-363. doi:10.1016/j.paid.2004.04.013

- Howard, R. W. (2006). A complete database of international chess players and chess performance ratings for varied longitudinal studies. *Behavior Research Methods*, 38(4), 698-703.
- Howard, R. W. (2008). Linking extreme precocity and adult eminence: A study of eight prodigies at international chess. *High Ability Studies*, 19(2), 117-130.
doi:10.1080/13598130802503991
- Howard R. W. (2009). Individual differences in expertise development over decades in a complex intellectual domain. *Memory & Cognition*, 37(2), 194-209.
doi:10.3758/MC.37.2.194
- Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth. *Behavioral & Brain Sciences*, 21, 399-442. doi:10.1017/S0140525X9800123X
- Jastrzemski, T. S., Charness, N., & Vasyukova, C. (2006). Expertise and age effects on knowledge activation in chess. *Psychology and Aging*, 21(2), 401-405.
doi:10.1037/0882-7974.21.2.401
- Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, 125(4), 331-359. doi:10.1037/0096-3445.125.4.331
- Maass, A., D'Ettole, C., & Cadinu, M. (2008). Checkmate? The role of gender stereotypes in the ultimate intellectual sport. *European Journal of Social Psychology*, 38(2), 231-245. doi:10.1002/ejsp.440

Masunaga, H., & Horn, J. (2001). Expertise and age-related changes in components of intelligence. *Psychology and Aging, 16*(2), 293-311. doi:10.1037/0882-7974.16.2.293

Medic, N., Starkes, J. L., & Young, B. W. (2007). Examining relative age effects on performance achievement and participation rates in Masters athletes. *Journal of Sports Sciences, 25*(12), 1377-1384. doi:10.1080/02640410601110128

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *The Psychological Review, 63*, 81-97. doi:10.1037/h0043158

Morrow, D. G., Miller, L. M. S., Ridolfo, H. E., Magnor, C., Fischer, U. M., Kokayeff, N. K., & Stine-Morrow, E. A. L. (2009). Expertise and age differences in pilot decision making. *Aging, Neuropsychology, and Cognition, 16*(1), 33-55. doi:10.1080/13825580802195641

Moul, C. C., & Nye, J. V. C. (2009). Did the soviets collude? A statistical analysis of championship chess 1940–1978. *Journal of Economic Behavior & Organization, 70*(1-2), 10-21. doi:10.1016/j.jebo.2009.01.009

Nunes, A., & Kramer, A. F. (2009). Experience-based mitigation of age-related performance declines: Evidence from air traffic control. *Journal of Experimental Psychology: Applied, 15*(1), 12-24. doi:10.1037/a0014947

- Puddephatt, A. J. (2008). Incorporating ritual into greedy institution theory: The case of devotion in amateur chess. *The Sociological Quarterly*, 49(1), 155-180.
- Puddephatt, A. J. (2003). Chess playing as strategic activity. *Symbolic Interaction*, 26(2), 263-284. doi:10.1525/si.2003.26.2.263
- Reingold, E. M., Charness, N., Schultetus, R. S., & Stampe, D. M. (2001). Perceptual automaticity in expert chess players: Parallel encoding of chess relations. *Psychonomic Bulletin & Review*, 8(3), 504-510.
- Riegel, K. F., & Riegel, R. M. (1972). Development, drop, and death. *Developmental Psychology*, 6(2), 306-319. doi:10.1037/h0032104
- Riegel, K. F., Riegel, R. M., & Meyer, G. (1967). A study of the dropout rates in longitudinal research on aging and the prediction of death. *Journal of Personality and Social Psychology*, 5(3), 342-348. doi:10.1037/h0024298
- Roring, R. W., & Charness, N. (2007). A multilevel model analysis of expertise in chess across the life span. *Psychology and Aging*, 22(2), 291-299. doi:10.1037/0882-7974.22.2.291
- Rubin, E. (1960). The age factor in Master chess. *The American Statistician*, 14(5), 19-21.
- Salthouse, T. A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology*, 113, 345-371.

- Schulz, R., & Curnow, C. (1988). Peak performance and age among superathletes: Track and field, swimming, baseball, tennis, and golf. *Journal of Gerontology*, *43*, 113-120.
- Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, *61*(4), 394-403.
- Stones, M. J., & Kozma, A. (1984) Longitudinal trends in track and field performances. *Experimental Aging Research* *10*, 107-110. doi:10.1080/03610738408258552
- Tabachnick, B. G., & Fidell, L. S. (2007). Using Multivariate Statistics: 5th edition. Pearson Education Inc., Boston.
- Twisk, J.W.R. (2006). Applied Multilevel Analysis: A Practical Guide. Cambridge University Press, UK.
- Vallerand, R. J., Mageau, G. A., Elliot, A. J., Dumais, A., Demers, M., & Rousseau, F. (2008). Passion and performance attainment in sport. *Psychology of Sport and Exercise*, *9*(3), 373-392. doi:10.1016/j.psychsport.2007.05.003
- Vollstädt-Klein, S., Grimm, O., Kirsch, P., & Bilalić, M. (2010). Personality of elite male and female chess players and its relation to chess skill. *Learning and Individual Difference*. doi:10.1016/j.lindif.2010.04.005
- Waters, A. J., Gobet, F., & Leyden, G. (2002). Visuo-spatial abilities in chess players. *British Journal of Psychology*, *93*, 557-565.

- Weir, P.L., Kerr, T., Hodges, N.J., McKay, S.M., & Starkes, J.L. (2002). Master swimmers: How are they different from younger elite swimmers? An examination of practice and performance patterns. *Journal of Aging and Physical Activity, 10*, 41-63.
- Young, B. W., Medic, N., Weir, P. L., & Starkes, J. L. (2008). Explaining performance in elite middle-aged runners: Contributions from age and from ongoing and past training factors. *Journal of Sport & Exercise Psychology, 30*(6), 737-754.
- Young, B. W., & Starkes, J. L. (2005). Career-span analyses of track performance: Longitudinal data present a more optimistic view of age-related performance decline. *Experimental Aging Research, 31*(1), 69-90. doi:10.1080/03610730590882855
- Young, B. W., Weir, P. L., Starkes, J. L., & Medic, N. (2008). Does lifelong training temper age-related decline in sport performance? Interpreting differences between cross-sectional and longitudinal data. *Experimental Aging Research, 34*(1), 27-48. doi:10.1080/03610730701761924