Is Age Really Cruel to Experts? Compensatory Effects of Activity

Nemanja Vaci, Bartosz Gula, and Merim Bilalić
Alps Adria University Klagenfurt

Age-related decline may not be as pronounced in complex activities as it is in basic cognitive processes, but ability deterioration with age is difficult to deny. However, studies disagree on whether age is kinder to more able people than it is to their less able peers. In this article, we investigated the "age is kinder to the more able" hypothesis by using a chess database that contains activity records for both beginners and world-class players. The descriptive data suggested that the skill function across age captures the 3 phases as described in Simonton's model of career trajectories: initial rise to the peak of performance, postpeak decline, and eventual stabilization of decline. We therefore modeled the data with a linear mixed-effect model using the cubic function that captures 3 phases. The results show that age may be kind to the more able in a subtler manner than has previously been assumed. After reaching the peak at around 38 years, the more able players deteriorated more quickly. Their decline, however, started to slow down at around 52 years, earlier than for less able players (57 years). Both the decline and its stabilization were significantly influenced by activity. The more players engaged in playing tournaments, the less they declined and the earlier they started to stabilize. The best experts may not be immune to aging, but their previously acquired expertise and current activity enable them to maintain high levels of skill even at an advanced age.

Keywords: aging, expertise, skill acquisition, linear mixed-effect modeling, chess

Supplemental materials: http://dx.doi.org/10.1037/pag0000056.supp

Over the years, memory and problem-solving abilities decline (Deary et al., 2009; Naveh-Benjamin & Old, 2008), while reaction times (RTs) slow down by almost two thirds when we compare 60-year-olds with 20-year-olds (Salthouse, 1984). Increased age is also negatively associated with learning, spatial ability, reasoning, and motor skills (Lindenberger & Baltes, 1994; Salthouse, 1991; Verhaeghen & Salthouse, 1997). The conclusion from a large number of studies is that the aging process negatively influences cognitive processes (Salthouse, 2010). On the other hand, the age-related decline observed in laboratory settings is often different from the age-related change in real life skills. For example, job performance reveals little or no decline through the years (Salthouse & Maurer, 1996; see also, Ng & Feldman, 2008), while in complex jobs, it is actually the case that age-related improvement is possible (Sturman, 2003). Evidence from meta-analyses (Waldman & Avolio, 1993) shows that job performance ratings have a stronger positive relation with age for professionals compared to nonprofessionals. This pattern of results is more in line with the common belief that older people are also wiser (but see Redzanowski & Glück, 2013; Staudinger & Glück, 2011). After all, senior people regularly hold positions of leadership and responsibility within universities, businesses, and governments (Horn & Masunaga, 2006; Horton, Baker, & Schorer, 2008). In short, studies on real life skills show that domain-related knowledge and experience could present a promising way of compensating for the age-related decline in cognitive processes. Here we investigate the onset and patterns of age-related decline in chess, a game that combines various higher and lower level cognitive processes. The main goal of the present study is to examine whether expertise, largely based on accumulated domain-specific knowledge (Chase & Simon, 1973; Gobet & Simon, 1996), and expertise-related activity can offer a way to compensate for age-related decline in a complex cognitive skill such as chess.

Age Is Kinder to the More Able

The research on the effects of age on ability has a long tradition. One of the first investigations was conducted by Thorndike, Bregman, Tilton, and Woodyard (1928), who investigated age-related decline in general intelligence (Army Alpha) as a function of initial ability level or as they called it "innate ability" (p. 17). They concluded that "it is probable that the influence is very slight, that the ablest man and the ordinary man show nearly the same curve." In contrast, Blum and Jarvik (1974) showed stronger effects of initial ability on age-related decline, and coined the "age is kinder to the initially more able" hypothesis. In their study, more able people showed a lesser decline by age on tests in cognitive battery than ordinary people did.

The analogous idea was tested in different domains of expertise with skill levels as a moderator of age-related decline. For example, expertise studies of complex problem solving skills such as chess provide some evidence for this hypothesis. Roring and Charness (2007, see also Almuhtadi, 2011) observed that rating scores of chess players decline over the years. Most importantly, higher ranked (and therefore more able or capable) players decline less than their lower-ranked colleagues in later years. Masunaga and Horn (2001) showed that increase in expertise of players (in the case of the board game GO) reduces or even stops the age-related decline in deductive reasoning and working memory. Ramscar, Hendrix, Shaoul, Milin, and Baayen (2014) used statistical and computational models to tackle the question of decline in processing speed for language comprehension. They proposed that the decline in the processing speed of older adults reflects expertise in language, that is, increased knowledge and vocabulary size, rather than cognitive decline per se.

On the other hand, Salthouse (2010) in his seminal overview of cognitive decline showed that ability does not interact with decline patterns. In one of the studies, Salthouse divided participants based on their performance and compared skill trajectories. For cognitive tasks such as word recall, spatial relation processing, and digit recall, all participants declined with age. However, there was essentially no difference between the best, average, and worst performing participants in the rate of decline. Similarly, Wilson et al. (2009) showed that the steepness of cognitive decline in people is equal across levels of education. Other studies have found the opposite effect, that is, age is kinder to the initially more average. In chess and bridge, for example, older players derived less benefit from increased skill when remembering briefly presented game situations (Charness, 1979, 1981). The results indicate that the acquired knowledge is not protective for the domain-related tasks that rely on the processes undergoing normative age-related decline such as memory and speed of processing (see the computational model in Mireles & Charness, 2002). Similarly, initially more prolific scientists have steeper age-related decline, as measured by the number of publications, after their peak than their less productive colleagues (Horner et al., 1986).[1]

Model of Career Trajectories and Landmarks

The opposite pattern of results could be accounted for by Simonton's model of career trajectories and landmarks (Simonton, 1977, also Simonton, 2015). This model was devised to explain creative potential and output throughout a person's life. It assumes a two-step process, where in the first step the creator begins with a supply of latent creative potential that is being transformed into creative ideas. In the second step, the ideas are being translated into products (see Simonton, 1983, p. 79). Simonton's model of career trajectories states that productive individuals with higher output show a steeper decline after their peak, which is the reverse of the hypothesis that age is kinder to the initially more able. According to the model, postpeak decrease will be proportional to the prepeak increase.

The model results in three different stages of creative process throughout a person's lifetime. At the beginning of their career, the creator quickly generates new ideas, which are produced more quickly than they are converted to final products. The fast production of ideas is characterized as the increase of the age related function in the earlier period of life. Productivity later levels off at the peak of the function, usually occurring around the 30s or 40s. After the increase and peak, the conversion of ideas to products is becoming faster than generation of new ideas, resulting in a postpeak decline in production of ideas. Productivity starts to decline at a steady rate, which is influenced by people's prepeak increase. In the case of more productive people, creativity tends to rise rapidly to the peak and decline just as rapidly after the peak is reached. In contrast, the creativity of less productive people approaches the peak more gradually, and declines very slowly (Simonton, 2010). The higher the productivity, the more quickly creative potential is worked into contributions (Simonton, 1983). Finally, the decline starts to stabilize at a particular point and slowly approaches the horizontal axis (Simonton, 1988). This third phase represents the change of the postpeak decline. The change is defined by an inflection point of the age related function, where the steepest decline from the peak begins to stabilize.

Studies in the domain of creative productivity showed support for this type of age-related curve (Dennis, 1966; Gingras, Larivière, Macaluso, & Robitaille, 2008; Lehman, 1953; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Simonton, 2015; Zuckerman, 1977). However, most of the studies in other domains of expertise do not establish the appropriate skill trajectories for their data (e.g., Howard, 2005, 2014a, 2014b; Roring & Charness, 2007; Salthouse, 2010). The common approach is to apply a certain order of polynomial in age to the dataset, as these polynomial effects tend to explain a lot of variance, without any procedure of model selection and comparative goodness of fit testing (Simonton, 1988, 1997). The main problem with this practice is that outcomes are inconclusive and differ between studies, as the effects obtained in the linear model depend heavily on the theoretical function that was applied to the data. Here we employ the model comparison procedures on the chess data to examine the optimal theoretical function that can explain age-related function.

Age Effects in Chess

Chess is a particularly suitable domain for tackling the question about expertise and age-related declines for a number of reasons. First, it is a complex cognitive domain that encompasses numerous cognitive processes such as perception (Bilalić, Kiesel, Pohl, Erb, & Grodd, 2011; Rennig, Bilalić, Huberle, Karnath, & Himmelbach, 2013), attention (Charness, Reingold, Pomplun, & Stampe, 2001), memory (de Groot, Gobet, & Jongman, 1996; Gong, Ericsson, & Moxley, 2015), and problem solving (Bilalić & McLeod, 2014; Connors, Burns, & Campitelli, 2011; Moxley, Ericsson, Charness, & Krampe, 2012). Second, it provides an objective and reliable measure of skill (Elo, 1978). Most importantly, there are a number of datasets that enable one to follow chess players' performances throughout their lives (Bilalić et al., 2009; Chabris & Glickman, 2006; Charness & Gerchak, 1996; Howard, 2005). These databases have been used to tackle questions such as gender difference (Bilalić et al., 2009; Chabris &

---

[1] The particular effect requires a closer examination as it runs counter to a large number of studies where researchers find cumulative advantage of a better start in science (see Allison, Long, & Krauze, 1982). In other words, researchers that are more prolific exhibit smaller postpeak decline in publication output compared with less prolific colleagues.

Glickman, 2006; Charness & Gerchak, 1996; Howard, 2008, 2009; Knapp, 2010) and skill acquisition trajectories (Gaschler et al., 2014; Howard, 2014a).

The chess databases also offer unbiased and reliable information about the tournament activity level of practitioners, something that is often missing in other domains. A common finding in previous skill acquisition studies is that continued expertise-related activity is necessary to sustain a high level of performance (Ericsson, Krampe, & Tesch-Römer, 1993). This is particularly the case in the research on aging and expertise, because previous studies show that expertise-related activity serves to counterbalance for age-related decline in performance (Bedard, 1989; Charness, Krampe, & Mayr, 1996; Ericsson, 2004; Krampe & Ericsson, 1996). The chess datasets offer information about the tournament activity of players and can therefore be used to tackle the interaction between age and expertise-related activities.

The most relevant study for our current research was conducted by Roring and Charness (2007), who applied multilevel modeling on a chess database. They showed that chess players reach their peak performance at around 43 years and that there is a dropdown of scores for all players after the peak. Most importantly, Roring and Charness demonstrated that the higher ranked players decline less in later years compared with lower ranked players. They also showed that the number of tournament games played, an expertise-related activity (Campitelli & Gobet, 2008, 2010), does not interact with age-related trajectories.

Roring and Charness (2007) modeled the data using a quadratic function, which might not be the best choice if we are interested in the third phase of decline stabilization in Simonton's model. More importantly, they used the international chess database (FIDE database), which suffers from serious methodological problems (Vaci, Gula, & Bilalić, 2014). The FIDE database only provides records for the very best practitioners and completely excludes weak players (and, to a lesser extent, average players too), as

illustrated in Figure 1A. Furthermore, the FIDE database has restriction of values for tournament activity, as can be seen in Figure 1B. In essence, Roring and Charness could only differentiate between highly skilled chess players whose expertise-related activity may not be logged accurately. The very best chess players may decline more slowly than their slightly less able colleagues, but it is still unclear what happens with average players.

Current Study

The main goal of the present study was to examine age-related decline in a complex cognitive skill. We were particularly interested in the hypothesis that decline is positively altered by the overall ability of practitioners. The previous studies (e.g., Almuhtadi, 2011; Howard, 2012; Roring & Charness, 2007) produced inconclusive results because they suffered from methodological problems, such as the use of inappropriate functions to describe skill trajectories (Simonton, 1997), lack of activity records (Ericsson & Moxley, 2012), and range restriction of the ability measure (Vaci et al., 2014, see Figure 1). As we will show in this study, the restriction of range for rating scores and tournament activity has serious consequences for the conclusions. We circumvent these problems by using the German chess database (Bilalić et al., 2009), which encompasses the full range of ability and provides reliable records of expertise-related activity (tournament activity).

Our general assumption is that the increase to the peak, and the subsequent decline, could follow Simonton's model of career trajectories and landmarks (Simonton, 1997). We assume that cubic functions provide a better description of both the German and FIDE chess databases because they capture all three phases (peak, decline, and beginning of stabilization) of Simonton's model (see the Fitting the Curve subsection in Results). We also expect the decline after the peak to be proportional to the increase
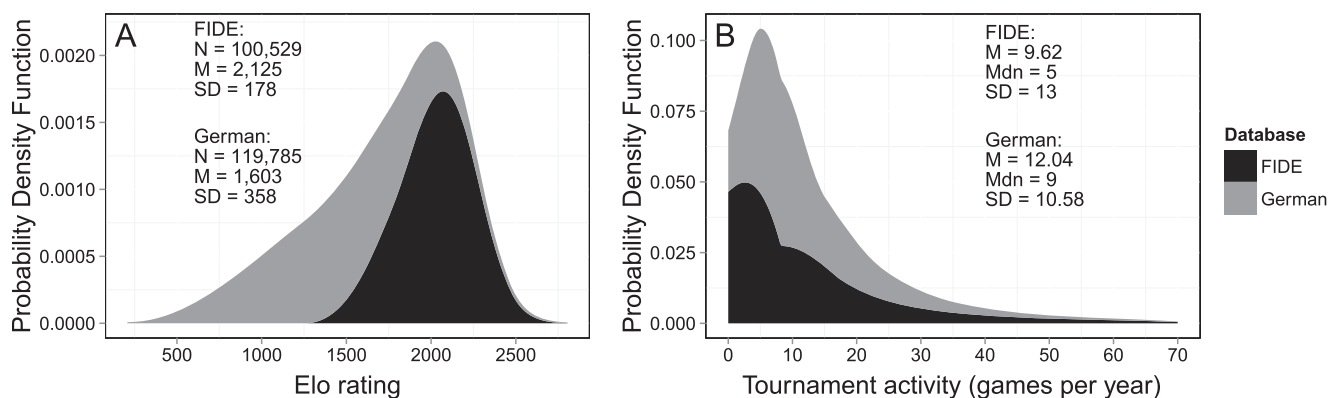

Figure 1. Probability density distribution of chess skill and tournament activity. (A) Probability density distribution of chess skill as measured by Elo rating in FIDE (dark gray) and German database (light gray). The datasets contain a similar number of players, but they differ in the shape of distribution and coverage. The only overlap between them is at the highest values of the German database and the lowest values of the FIDE database. Average players have around 1,500 rating points, while experts are considered to be players who have more than 2,000 rating points (see Method section for a detailed explanation). The Y-axis is the probability of Elo points across all players (density). (B) Probability density distribution for tournament activity as measured by number of played games for every player in a year, in FIDE (dark gray color) and German database (light gray color). The distributions of tournament activity overlap, but the German dataset logs more records compared to the FIDE dataset.

before the peak. If players improve at a faster rate before the peak then they should also decline more quickly. Conversely, players with a shallow increase to the peak should also have a shallow decrease after the peak (see the Age is Kinder to the Initially More Able subsection in Results).

We are also interested in how the two different datasets with full (German) and restricted range (FIDE) differ in the case of age-related declines. As has been shown, analyses on the restricted FIDE dataset indicate that the very best players in the world have a smaller decline in later years compared with slightly worse but still very good players (Almuhtadi, 2011; Roring & Charness, 2007). If age is indeed kinder to the initially more able, then we would also expect nonexpert players to decline more quickly than experts in the German database, where there are no restrictions in skill range.

Extending the previous results, we analyze the skill acquisition function in chess expertise for the third phase in Simonton's model, namely stabilization of decline in later years (Dennis, 1966; Lehman, 1953; Zuckerman, 1977). If age is kinder to the initially more able, then we would expect the stabilization point to be reached earlier in experts than in nonexpert players (see the Stabilization of Age-Dependent Decline subsection in Results). Given the importance of expertise-related activities in later years (Ericsson, 2004; Krampe & Ericsson, 1996), we also expect a preserving effect of tournament activity on skill (see the Modeling of Age and Activity Effects subsection in Results). In other words, decline in experts will be mitigated by activity, as evidenced by an earlier stabilization point in players who practice more.

Method

Chess Skill

Chess skill is measured on a continuous scale, which reflects the performance of player against player. The Elo rating, named after Arpad Elo who introduced the scale as a measure of chess skill (Elo, 1978), is applied in a similar way across the world. The Elo rating is inferred from player versus player outcomes and uses normal distribution to measure performance. The theoretical mean is set at 1,500 Elo points and theoretical standard deviation is 200 points. Multiple groups of players can be identified within the distribution of the Elo scale. For example, beginners have around 800 Elo points, average players around 1,500, masters above 2,200 and the very best players, grandmasters, around 2,500 Elo points. Players above 2,000 points are usually considered as experts.

Databases

We used two large databases, one maintained by the International Chess Federation (FIDE; see Howard, 2006 for more information) and the other by the German Chess Federation (DSB: Deutscher Schachbund; see Bilalić et al., 2009 for more information). The FIDE database collects records from the 1970s to 2010, which enables researchers to estimate trajectories over the course of life. The FIDE rating list includes only the best players in the world. In the beginning, FIDE kept only the records of players above 2,200 Elo points. The entry limit has subsequently been lowered, but the database still contains only the best players and some above-average players, as depicted in Figure 1A. Besides the restriction in skill range, the FIDE database also puts a restriction

on the games played. The tournaments are only recorded and logged in the FIDE database if they are registered as FIDE events (for a considerable fee). Many national federations do not possess the necessary means to register their tournaments with FIDE. Consequently, only a fraction of the games played by their players is captured by the FIDE database. The FIDE database provides an imposing amount of data for the best chess players in the world, but it unfortunately imposes severe restrictions on skill range and playing activity (see Figure 1A and 1B).

The German database represents one of the biggest and best-organized national chess databases in the world. All German tournaments are rated, including club championships that feature not only competitive players but also hobbyists. The German database collects records from 1980 to 2007, and keeps records of FIDE tournaments organized in Germany. This results in a reliable estimation of the activity of players and a full range of skill from beginners to the world's best players, as depicted in Figure 1A and 1B.

Both databases have encountered criteria changes for the data keeping throughout their history (see, e.g., Howard, 2006). As we have shown in the cohort analysis (see additional supplemental material for further analyses on both datasets), these changes do not seem to have influenced the skill trajectories across the players, at least not in the German database we use for the conclusion. More importantly, the rating employed in the German database is based on the same assumptions as the Elo rating. The two databases measure the same skill, as is shown by the high correlation between FIDE and German ratings in the players who both possess both German and FIDE ratings ($r = .93$ in Bilalić et al., 2009).

Preliminary Data Screening

Preliminary data screening was performed to ensure that potential data entry mistakes were accounted for. We followed the same procedure as in previous studies (Dennis, 1966; Roring & Charness, 2007; Simonton, 1977) and excluded players under 10 and over 80 years. Additionally, players with only one observation were excluded to ensure that analyses were performed for active players only, as at least two data points are required to make inferences about the change of rating scores.

The FIDE database used here includes all records up to 2010 and after the preliminary data screening, it comprised 2,916,227 observations with 100,529 players for the FIDE database (with an average of 42 observations per player). The German database included the data up to April 2007 and encompassed 2,072,176 observations with 119,785 individuals (with an average of 33 observations per player). We provide the descriptive statistics for the main variables used in the models (also collapsed over the participants' age) in Table A1 and A2 in Appendix A.

Overview and Rationale for the Analyses

The databases were analyzed in several iterations (datasets and the R code for the main and additional analyses can be found in supplemental materials). In the first step, we investigated the descriptive patterns of age-related effects on rating scores. To examine whether the two databases differ in the descriptive patterns, we applied the local polynomial regression in R statistical environment (R Core Team, 2013) and compared data-driven

trajectories of rating as a function of age. More importantly, we assessed the best theoretical function behind chess skill acquisition. This was obtained by comparing models with different polynomial terms on age as a predictor variable, ranging from linear to the seventh polynomial.

In the second step, mixed-effect models were applied on both databases using the *lme4* package in R (see Baayen, Davidson, & Bates, 2008; Baayen & Milin, 2010; Bates, 2005; Radanović & Vaci, 2013; for details about the *lme4* application). The main idea behind this statistical approach is to estimate and control additional sources of variability in the dependent variable, which are not influenced by the factors of interest (fixed effects). The most frequent spurious variable is the subject factor, where individual differences in genetic, developmental, environmental, social or chance factors are influencing variability in the dependent variable. More importantly, the repeated measurement data, as in the case of this study, is hierarchically organized. Multiple observations measured for each player are correlated and clustered within a player. One way to account for the additional variance and serial correlation is to treat players as random effects in the model. By doing this, one is estimating skill acquisition curves for each player separately. Another interesting possibility in the case of linear mixed-effect modeling is the option to examine the relation between fixed factors and random effects in more detail.

Similarly to Roring and Charness (2007), we empirically tested whether linear, quadratic, and cubic effects of age, as well as the number of played games per tournament, predict chess skill. All these factors were included as fixed effects. Additionally, to better investigate the effect of expertise-related activity and decline of rating scores due to periods of the inactivity in players' competitive careers, we calculated the time difference between logged tournaments for each player. We call this inactivity "stale play." The predictor "stale play" illustrates the time span in rated chess activity and by adding it to the model we are adjusting the function for the time of inactivity for each player. We also tested whether the adjustment of the intercept and slopes was necessary for players and the fixed effects in the model. Comparison of models showed that random effect of intercept for players and slope for linear effect of age added to the goodness of fit (see Appendix B). To illustrate how this modeling is applied, we will specify the main model used in our study in Equation 1. In the model, *y* refers to the rating of an individual player (*p*) at a time point (*i*). Intercept of the model is decomposed into two parts: a global intercept (β_O) and adjustments of intercept for each player (P_{0p}). The linear effect of age is adjusted in the same manner. A single global estimate for the linear effect of age (β_1) and the degree to which the linear effect of age for each player deviates from the global estimate (P_{1p}), while allowing correlation between intercept and slope deviations. We did not impose zero correlation between random intercepts and slopes in order to examine the relation between starting points and increase to the peak in more detail. All remaining terms are fixed effects and are shown in the common GLM notation.

$$y_{pi} = \beta_0 + P_{0p} + (\beta_1 + P_{1p})Age_i + \beta_2 Age_i^2 + \beta_3 Age_i^3 + \beta_4 Games_i$$
$$+ \beta_5 StalePlay_i + \beta_6 Age_i Games_i + \beta_7 Age_i^2 Games_i$$
$$+ \beta_8 Age_i^3 Games_i + e_{pi} \tag{1}$$

In the last step, we investigated the "age is kinder to the initially more able" hypothesis. This was examined by including a new factor in the model, which coded all players as either experts or nonexperts. By adding a new factor, we adjusted the estimated coefficients in the model and plotted life span trajectories for both kinds of player. A similar approach has already been used by Salthouse (2010) to examine the "age is kinder to the initially more able" hypothesis.

To sum up, we present first the descriptive and model-based results on the relationship between age and Elo ratings in the FIDE and German data, separately. We then show how expertise-related activity affects Elo ratings in general, age at peak, and the shape of postpeak decline, by adding the number of tournament games per rating period to the model. In the next step, we add an initial ability level factor (best vs. average players) to the model in order to test the "age is kinder to the initially more able" hypothesis. Finally, we calculate the inflection point of decline for both datasets and investigate whether this onset of stabilization phase is dependent on tournament activity, that is, the activity level of players.

Results

Pattern of Skill Acquisition in Chess

Descriptive pattern of skill trajectories. Data-driven, descriptive patterns of skill trajectories were obtained by fitting local polynomial regression—loess on averaged values per year (Cleveland, Grosse, & Shyu, 1992), for the FIDE and German data separately. The averaged Elo values (see Figure 2) were used as the dependent variables, while the age of participants was the predictor. The results show different skill acquisition trajectories in the two datasets (as partially discussed in Vaci et al., 2014). The German players show a steeper increase to the peak compared with the FIDE players (see Figure 2). They are starting as beginners at around 1,000 Elo points and tend to improve more in the first years of playing. In contrast, the FIDE players are already entering the dataset as highly skilled players (around 1,800 Elo points) and therefore show smaller effects of increasing skill prior to the peak. The peak performance is at approximately 39 years in both datasets. As expected, FIDE players are better on average than German players when comparing rating values.
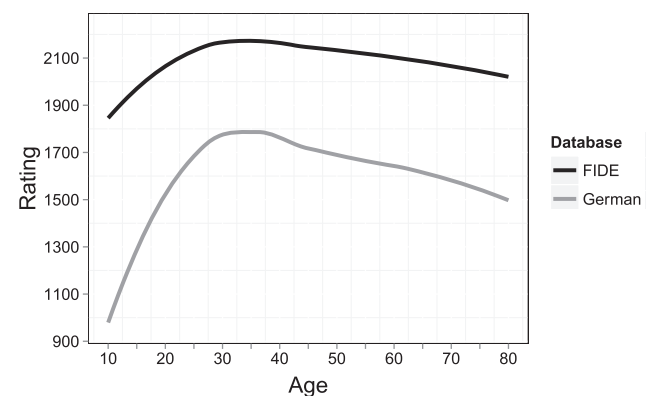

Figure 2. Data-driven age-related trajectories for averaged Elo scores, in the case of the FIDE (black) and German databases (gray).
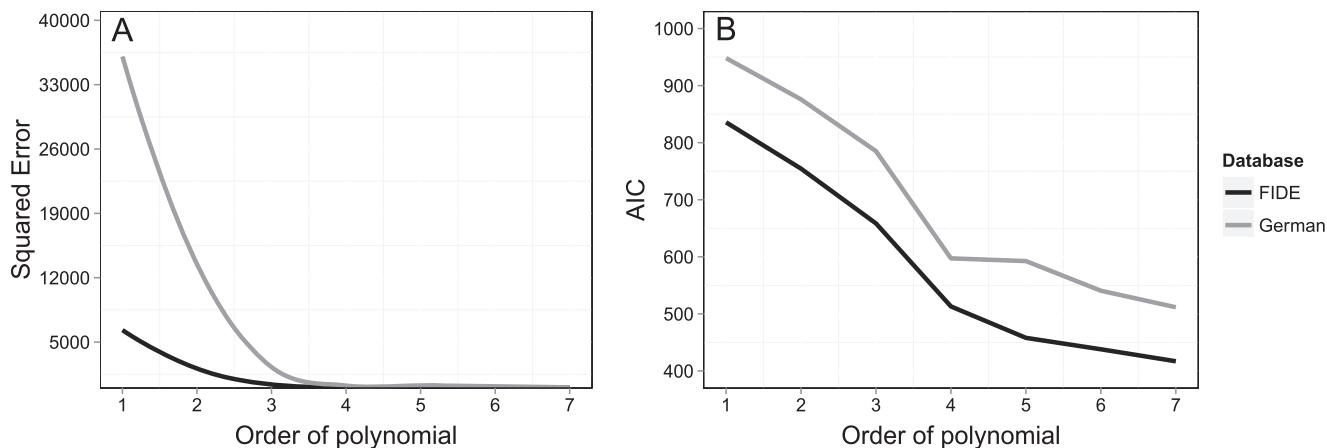
Figure 3. Examination of theoretical function. (A) Squared error comparison for different order of polynomials in age, calculated based on leave-one-out cross-validation, for FIDE database (black) and German database (gray). (B) AIC comparison for different order of polynomials in age, for FIDE and German databases.

Fitting the curve. The question remains as to what is the best theoretical function that can explain the behavior of the rating scores over the years. The most frequently used theoretical function in previous studies is the quadratic function, while the model proposed by Simonton (1997) assumes more complex behavior of age-related function, which could be approximated with the third polynomial or cubic function. The cubic function may also help to capture the stabilization of the decline in later years as obtained in the descriptive analyses (see Figure 2), something the quadratic function cannot capture.

To assess what is the most parsimonious order of polynomial for a given dataset, we used "leave one out cross-validation" (Browne, 2000; Roberts & Pashler, 2000). That is, parameters for all models were estimated on a training sample, which excluded one data point. After this, the estimated model was used to predict this particular data point. The fit of the model was examined by comparing AIC values and squared errors, summed across 80 simulation trials, for each model. Contrary to the fit of the function, AIC penalizes increasing number of parameters, thus, it discourages overfitting of the data. An important aspect when choosing the function is its plausible interpretation. That is to say, due to the vast amount of information gathered in datasets, adding more complexity tends to improve fit of the model automatically. By taking into account each of the two measures of fit and the interpretability of the function, we can determine the best and most parsimonious model. Because cross-validation is computationally demanding, we fitted linear regression with rating as the dependent variable and age of participants as the predictor, while we tested for the maximum of seventh polynomial. The analysis was performed on both the FIDE and German databases.

Both measures of goodness of fit show similar trends. The squared error of the model declines by adding more complexity, while the inflection of the squared error decline is around third order polynomial (see Figure 3A). The AIC values also show that the quality of the model also improves by adding more complexity. However, the inflection of AIC values occurs at fourth order of polynomial (see Figure 3B). If we take into account previous

descriptive analysis of the data (see Figure 2), third order of polynomial should offer a good fit to the data.[2]

Third order of polynomial is also justified based on the results obtained from the linear model (see Table 1). In both datasets, there is linear improvement in the beginning of the career illustrated as positive linear effect. The quadratic effect has a negative sign that indicates an inverted-U function, while positive cubic effects flattens the decline from the peak in later years. The results illustrate three different stages of expertise development. The linear and quadratic effects illustrate skill acquisition in the beginning of career and skill decline in later years, respectively. The cubic effect captures the right tail of the age-rating trajectories where players do not decline especially rapidly, or the decline even tends to stagnate (see Figure 2). This is an important insight into the process of age-related decline that had not been captured by previous studies.

Mixed-Effect Modeling

Modeling of age and expertise-related activity effects. Next, we modeled the rating of players using linear mixed-effect modeling (Baayen, Davidson, & Bates, 2008; Bates, 2005). The linear, quadratic, and cubic effects of age, the number of played games (tournament activity) and time span between played tournaments ("stale play") were included as fixed-effects, while random effects were participants. We also attested for possible by-players random slope adjustments. Only the linear effect of age required additional adjustment of the slope, and so it was included as a random effect. Goodness of fit of the models was compared for each new term that was added to the model (see Appendix B).

The results for the FIDE database presented in Table 2 show a similar pattern of effects as in Roring and Charness (2007). Random intercepts for participants and adjustment of the slopes for the linear effect of age correlate significantly ($r = -.60$). This means

---

[2] Final models were also computed with fourth polynomial of age. However, the results for the main hypothesis, that age is kinder to the initially more able, remained the same.

Table 1
Estimated Coefficients, Standard Errors, t-Values and p-Values for the Linear Model With Third Order of Polynomial, for the FIDE and German Databases

Effect	FIDE database				German database			
	Estimate	*SE*	*t*-value	Pr(>|z|)	Estimate	*SE*	*t*-value	Pr(>|z|)
Intercept	1876	.52	3555.5	<.00	961.5	.69	1389.2	<.00
Age	27.95	.06	434.2	<.00	77.63	.10	776.1	<.00
Age^2	−.789	.002	−356.8	<.00	−2.190	.003	−587.0	<.00
Age^3	.006	.0002	283.7	<.00	.017	.0003	460.2	<.00

that players with higher scores at the beginning have a shallower increase to the peak, while lower starting players have a steeper increase. Players in the FIDE database peak around 34 years. The first polynomial or linear function of age has a positive effect on rating scores, while the second polynomial has a negative one. As previously argued, this illustrates the improvement at the beginning of the career and the decline in later years. The third polynomial or cubic function has a positive effect on rating scores. The cubic effect captures the start of the decline stabilization in the right tail of the skill function.

The numbers of played games also add new information to the model; as can be seen, all three effects of age were significantly affected by the activity. The interaction between the second polynomial of age and games is negative, indicating that there is a greater decline in later years depending on the number of played games. The model shows that people who play more peak higher, but they also decline faster than their peers who play less. It seems that the experience gathered through playing does not preserve their skill in later years. Instead, it actually negatively influences their skill. The "stale play" predictor has a negative effect on the rating scores. The longer the time lag between two tournaments, the more a player is going to decline.

A model with the same predictor terms produced different results in the German database (see Table 2). The players that start with higher rating scores at the beginning of their careers have a shallower increase to the peak—around 38 years (the correlation between intercepts and slope is $r = −.61$). As in the FIDE database, linear and cubic effects of age have a positive influence on the rating in the German database, while the quadratic effect and "stale play" predictor have a negative influence. The interaction of the quadratic effect of age and games has a positive effect on rating, while the linear effect enters in negative interaction.[3]

Figure 4 illustrates hypothetical skill acquisition trajectories that result from entering different numbers of games played into the estimated model equation. Depicted are the trajectories of players who play very often (30 games per year) and those who play rarely (three games per year). The more active players in the FIDE database peak higher and slightly earlier than the less active players, but they decline faster after the peak (see Figure 4A). The results obtained on the German database lead to a substantially different conclusion about the preserving effect of activity—chess experts with more games played peak higher, but they also decline more slowly, than their less active counterparts (see Figure 4B).

The "age is kinder to the initially more able" hypothesis. To investigate the hypothesis that age is kinder to the initially more able or capable, we divided players based on their rating scores. The initially more able or expert group consisted of players with a

peak rating of 2,000 Elo points or more, while all other players were defined as nonexpert players.

The mixed-effect model stated in Equation (1) was also used to test the "age is kinder to the initially more able" hypothesis, with the following difference. An expertise level factor indicating whether players fall within the best player group or the nonexpert one was added as a fixed effect to the model, as well as seven terms for all interactions with age and games. See Appendix C for parameter estimates.

The results show an opposing pattern of effects in the FIDE and the German database. In the FIDE database, the expert and nonexpert players are improving and becoming better at a similar rate. After the peak, the best players tend to decline less compared with lower ranked players (see Figure 5A). The results based on the decline after the peak support the "age is kinder to the initially more able" hypothesis. However, the best players in the German database have a steeper increase to the peak, or rather, a faster improving rate at the beginning of their careers. The effect of age after the peak in this instance is the converse of that seen with the FIDE database. The best players decline more in comparison with their less able peers (see Figure 5B). The results obtained on the German database support or resemble the assumption of proportionality between prepeak increase and postpeak decrease in Simonton's model of career and landmarks.[4]

We calculated the age at peak in rating scores for the best and average players in both datasets. Nonexpert players in the FIDE database peak around 34 years, while experts' peak is at 38 years.

---

[3] In order to investigate whether the negative effect of practice in the FIDE database may be influenced by the restriction of tournament activity and not the restriction of the range of rating scores, we conducted two additional analyses. In the first, we truncated the German distribution below 1,500 Elo points to make it comparable to the FIDE database when it comes to skill range. In the second analysis we identified the 13,487 players that are registered in both datasets (see Table A3 in Appendix A for descriptive statistics). The both analyses replicated the positive, preserving effect of activity in the German database, but not in the FIDE database (see Intersection of the Datasets analysis in supplemental materials). This suggests that the different signs of the estimated games parameters in both databases are unlikely to be due to skill restrictions, but rather to the activity restrictions.

[4] The FIDE database logs only players above 1,500 Elo, while the German database also includes players weaker than 1500 Elo. This means that our nonexpert category includes differently skilled players in the FIDE and German datasets. To investigate whether the differences between the datasets could be explained by the lack of below-average players (under 1,500 Elo) in the FIDE database, we matched the two datasets for rating level by excluding the players in German dataset below 1,500 Elo ratings (see Ability Factor analysis in supplemental materials). The results, however, remained the same as in the main text.

Table 2

Estimated Coefficients, Standard Errors, t-Values and p-Values From the Linear Mixed-Effect Model, Separately for the FIDE and German Database

Effect	FIDE database				German database			
	Estimate	SE	t-value	Pr(>|z|)	Estimate	SE	t-value	Pr(>|z|)
Intercept	1918	1.19	1611.6	<.00	597.1	2.24	266.4	<.00
Age	15.96	.06	240.0	<.00	97.3	.17	545.2	<.00
Age^2	−.421	1.75e-3	−240.6	<.00	−2.539	5.63e-3	−450.9	<.00
Age^3	.002	1.71e-5	146.6	<.00	.0186	5.11e-5	365.3	<.00
Games	.672	.016	58.0	<.00	2.55	.058	45.8	<.00
Stale play	−.278	.028	−9.9	<.05	−8.743	.084	−104.0	<.00
Age:Games	.015	1.52e-3	10.0	<.00	−.135	7.93e-3	−17.1	<.00
Age^2:Games	−.0015	5.58e-5	−27.3	<.00	.0026	2.95e-4	9.0	<.00
Age^3:Games	.000016	5.88e-7	28.3	<.00	−.00001	3.06e-6	−5.8	<.00

The pattern is the opposite in the German dataset: less able players peak around 42, whereas experts peak around 37 years. Similar to the previous analysis of activity, we checked the trajectories for best and nonexpert players who play three or 30 games. The results again show a positive, skill-preserving effect of activity on decline in the German database, but not in FIDE.

Stabilization of age-dependent decline. The descriptive and linear mixed-effects model analyses show that the inflection point of decline occurs in the right tail of the skill function. We computed the second derivative by using estimated coefficients from the mixed-effect model. The second derivative can be described as the function that reflects the change of the slope of any differentiable function. By setting the second derivative of the estimated model equation to zero and solving for the age of participants, we obtained the inflection point, which, in the present model, reflects the change of slope from decreasing to increasing. The inflection point reflects the intersection of highest decline rate after the peak and the onset of the stabilization phase of the age-related decline.

Computations on the linear mixed-effect model (see Table 2) show that the inflection point for players in the FIDE database occurs around 66 years, while for German players it occurs around 55 years. More importantly, the computations on the model with the ability factor (experts and nonexpert players) show that the inflection point is expertise-dependent. In both datasets, experts start to stabilize sooner compared with less able players. In the FIDE database, the inflection point is 61 years for experts and 99 years for nonexpert players. In the German database, it is 52 years for experts and 57 years for nonexpert players. As the standard deviation tends to increase in the tail of age-related function in the case of the FIDE database (see Table A2 in Appendix A), one should be cautious when interpreting estimates for this database.

The expertise-related activity also influences the inflection of decline. As in the previous analyses, we used two hypothetical groups of players who played often (30 games per year) and rarely (three games per year). In the case of more active players, stabilization of decline starts earlier, at 63 years compared with 66 years in the FIDE dataset. The same pattern is found in the German database: Experts start to stabilize at 54 years whereas average players begin slightly later at 55 years.[5]

Discussion

Age may be cruel to basic cognitive processes (Deary et al., 2009; Naveh-Benjamin & Old, 2008; Salthouse, 1984) but there

are indications that expertise and expertise-related activity in complex skills may help to prevent serious age-related decline (Sturman, 2003; Waldman & Avolio, 1993). However, the results are far from conclusive (Howard, 2012; Roring & Charness, 2007) and previous studies suffer from various methodological problems. Here we used the archived data in the domain of chess to circumvent the shortcomings of previous research, and to examine whether expertise and practice moderate age-related decline in complex cognitive skills.

The Shape of the Age Curve in Chess

We examined the descriptive patterns of the age-related trajectories in chess expertise. Initially, players in both datasets improve and progress, while in later years they slowly decline. The optimal age-related skill function in chess captures the phases of age-related behavior proposed by Simonton's (1997) model of career trajectories and landmarks. The fitted cubic function captures the initial process of improvement, decline after the peak, and finally, the start of the decline stabilization. Previous studies showed evidence for this type of skill curve (Dennis, 1966; Lehman, 1953; Simonton, 1988; Zuckerman, 1977).

It is important to state that we do not assume that chess expertise functions in the same manner as creativity. In the case of Simonton's model, expertise acquisition is assumed to have taken place before the onset of a person's career. This is not the case in chess expertise, where prepeak increase is usually interpreted as the skill acquisition phase. However, the typical phases in both domains bear a resemblance to each other and Simonton's model provide a useful framework for explaining chess trajectories across players' life span. Even the notion of individual differences in Simonton's model, where the productivity rates differ among disciplines, could be connected to our results. Instead of the productivity differences among (academic) disciplines, in chess we obtained different trajectories for differently skilled practitioners.

---

[5] When we use four differently skilled groups based on quartiles in the databases, instead of two, we get essentially the same results. The ability level influences the peak of the function and the stabilization of decline (see Ability Factor analysis in supplemental materials). The higher the level of ability, the sooner the peak and the sooner the inflection point is observed.
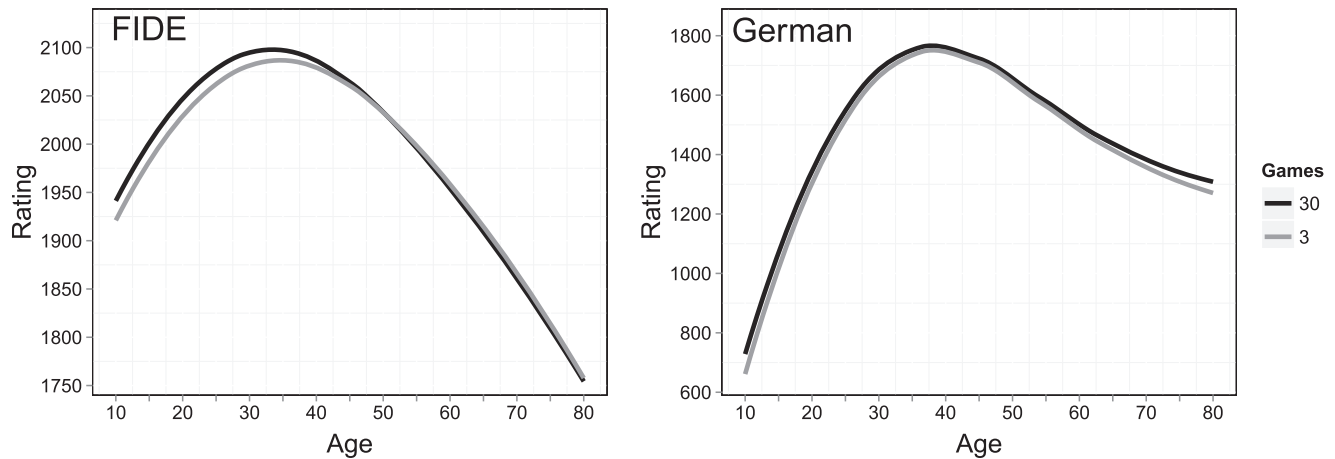
Figure 4. Effect of number of games played. (A) Skill acquisition in chess in the FIDE database, over years for two different theoretical players, first a player with three games played per year (black) and second a player with 30 games played per year (gray). (B) Skill acquisition in the German database, for a player with three games played per year and for a player with 30 games played per year. Please note that the range of the *y*-axis was not adjusted on the same range, as the raw Elo ratings are different between the databases.

Range Restriction of Ability and Activity

The results demonstrate the possible pitfalls of using the FIDE database. A database that misses a considerable number of players and does not record their activity is bound to produce unreliable results. The FIDE expertise curve is truncated because the players' initial skill acquisition process before reaching 1,500 Elo has not been recorded in the database. This results in a more prominent increase period for the German players, as they enter competitions mostly as beginners, whereas in the FIDE database the players are already experts when their scores first begin to be collected. These restrictions may provide a possible reason for other contradictory results in the literature. For example, studies with restricted FIDE

data regularly find differences between women and men in skill (Howard, 2005, 2006, 2014a but see, Bilalić and McLeod, 2006, 2007). The studies with unrestricted national databases, however, explain these differences through participation rates and dropout patterns (Bilalić et al., 2009; Chabris & Glickman, 2006; Knapp, 2010).

Similarly, the studies that use activity levels to estimate innate talent in chess (Howard, 2008, 2009) become difficult to interpret if a large number of games are missing from the records. The lack of reliable records of expertise-related activity is one of the reasons for the implausible results we found in the FIDE database where more tournament activity leads to greater postpeak skill decrease.
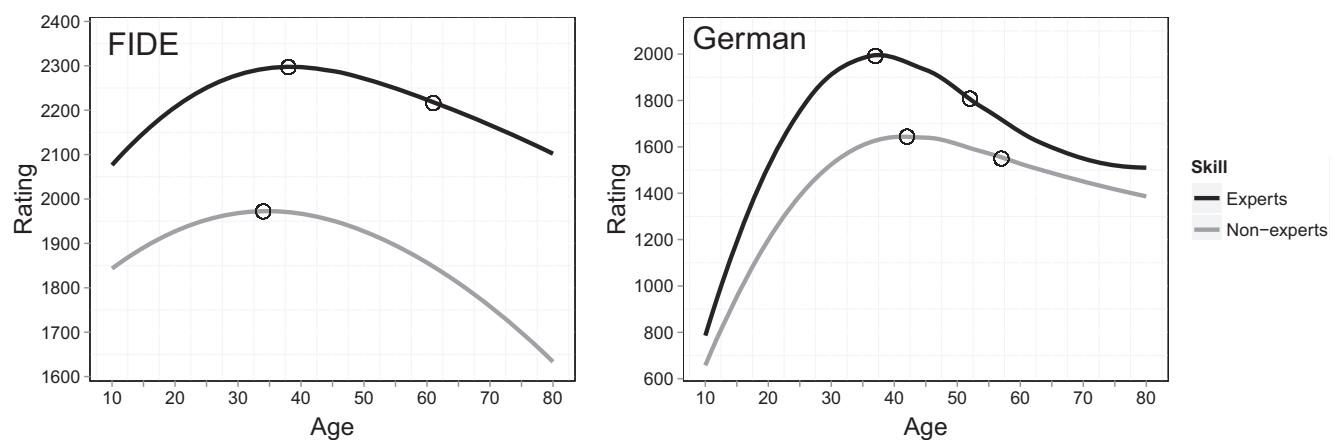
Figure 5. Life span trajectories for average and best chess players. (A) Age-related curves in the FIDE database for rating scores, measured in Elo points, for average players (black) and experts (gray). (B) Age-related curve in the German database for rating scores for average players and experts. The range of the *y*-axis was not adjusted on the same scale, as raw Elo ratings differ between datasets. The points represent maxima (peak values) and inflection points of decline (see stabilization of age-dependent decline), separately for experts and average players. The inflection point for average players in the FIDE is not shown because it is outside of the age range (99 years).

In the German database we find the opposite effect—tournament activity moderates the postpeak decline. One of the plausible reasons for the different effects of tournament activity in the two databases lies in the methodological constraints and logistical procedures (see Method, as well as Ericsson & Moxley, 2012). Another reason for this effect can be more motivational and environmental. In the later years, older players are encountering rapid rising, underrated younger players. The number of defeats against the younger opponents increases, producing larger drop-down of rating scores and discouragement, which may result in inactivity. Finally, playing more games allows better adjustments of the rating score than playing fewer games. In the case of older players that rarely participate, the rating score is going to undergo a smaller decline compared with older players that play more games.

The Effect of Activity on Aging in Chess

The activity levels of chess players played a role throughout their career. Players who were more active had steeper increases to the peaks and reached higher peaks than players who were less active. The activity also acted as an equalizer to the age-related decline observed after the peak. The players who were more active tended to decline more slowly than their less active colleagues and their decline also stabilized sooner. Additionally, time lag or "stale play" predictor negatively influences rating scores. This effect can be observed as complementary to the tournament play effect. On the one side, tournament play tells us how players decline or improve based on the expertise-related activity during rating periods. On the other side, the "stale play" illustrates how periods of inactivity and the time lapse between participation in chess tournaments influences overall skill decline. Both point out the major influence of expertise-related (in)activity on the chess skill development throughout the life span.

The results confirm previous studies, which demonstrated that activity plays an important role in the preservation of cognitive functions (Bedard, 1989; Ericsson, 2004; Krampe & Ericsson, 1996). It should be noted that we did not measure practice or even deliberate practice (Ericsson et al., 1993) but rather players' activity reflected in their tournament play. The number of games played in those tournaments does not constitute deliberate practice, which in chess has been defined as solitary study of books and focused tournament preparation (see Charness et al., 2005). While there are contradictory results on whether deliberate practice or tournament play is a better predictor (Campitelli & Gobet, 2008, 2011; Howard, 2012; Ericsson, & Moxley, 2012), it is beyond doubt that both factors play a huge role in chess skill. It is therefore conceivable that the inclusion of direct deliberate practice indicators (e.g., protocols of practice hours), beside the already incorporated tournament play, would result in even stronger preserving effects in cognitive decline.

Comparison of Age-Related Decline With Other Domains

Studies in gerontology and the aging approach showed that age-related decline is inevitable in most areas of cognitive functioning (Deary et al., 2009; Naveh-Benjamin & Old, 2008; Salthouse, 1984, 2010). One way to examine the magnitude of chess age-related decline is to calculate the effect size of the decline and compare it with previous studies in the area of cognitive age-related decline. Meyer et al. (2001) published an overview of estimated effect sizes or correlations for different type of cognitive processes and functioning. We adapted this approach and calculated the effect size for age-related decline in chess. Correlation of postpeak rating scores with age shows that this effect for the FIDE database is .19, while for the German database it is .22. These coefficients are much smaller than the age-related decline for episodic memory (.33), reasoning (.40), and speed of processing (.52). Similar results were found in a recent meta-analysis by Moxley and Charness (2013). The performance on the recall of briefly presented chess positions declines more with age ($r = -.49$) than the performance on the choice of the best move ($r = -.28$). These findings fit within the pattern where the age-related decline has been observed in basic cognitive processes (Horn & Masunaga, 2006; Lindenberger & Baltes, 1994; Salthouse, 1991), whereas the decline has been much lower or even absent in real life professions and complex cognitive skills (Salthouse & Maurer, 1996; Sturman, 2003).

Age Is in Long Term "Kinder to the Initially More Able"

The final step in our study was to examine the "age is kinder to the initially more able" hypothesis. The results obtained with linear mixed-effect modeling indicate that there are differences between the FIDE and German databases. Experts in the FIDE database decline less after the peak compared with nonexpert players. This effect is in line with the studies performed on the same dataset (see Roring & Charness, 2007) and supports the "age is kinder to the initially more able" hypothesis (Blum & Jarvik, 1974). The experts in the German database, on the other hand, have a steeper decline after the peak than their less able peers. Experts from this dataset display a steeper prepeak increase and postpeak decline than nonexpert players. These results do not completely favor the "age is kinder to the initially more able" hypothesis. Rather, they show evidence for Simonton's (1997) model of career trajectories and landmarks, as the prepeak increase is proportional to the postpeak decrease (Dennis, 1966; Lehman, 1953; Zuckerman, 1977).

One of the novelties of our study is a demonstration of previously unexplored effects of expertise and activity on age-related decline in complex cognitive skills. That is, we show that the postpeak decline starts to stabilize at one moment. This effect is in line with Simonton (1997), who showed that the postpeak curve becomes concave upward and approaches zero decline asymptotically. We estimated the onset of the decline stabilization and showed that it depends on expertise and activity. Experts display an earlier stabilization of decline compared with less able players, while activity moves the stabilization point to an earlier age. To our knowledge, this is the first study to illustrate the preserving effects of expertise and tournament activity on the tail of the age-dependent decline. If this effect can be replicated in future studies, it can give us another perspective on the preserving effects of expertise and activity. Previous work assumes a monotonic benefit of practice (Ericsson & Charness, 1994; Ericsson et al., 1993), namely, that improvement in performance is proportional to the current amount of activity. We replicated this effect and have

shown that postpeak practice offsets the age-related decline (see Figure 4). Beyond this, we have also demonstrated that previously accumulated practice has long-term effects on the stabilization phase. Experts, who most likely became experts through accumulation of practice, start to stabilize earlier (52 years) than nonexpert practitioners (57 years), who have probably accumulated much less practice in their lifetimes. These results also extend the implications of the "age is kinder to the initially more able" hypothesis. They not only show that the immediate decline after the peaks is important, but also that the tail of the distribution varies and is influenced by expertise and activity.

In summary, we show that performance declines as people get older, even in complex cognitive skills. The decline, however, is considerably smaller than in other basic cognitive processes. The results indicate that age may be crueler in the beginning of the decline to more able practitioners, but we also demonstrate the more subtle effects of expertise and activity. In line with Simonton's model (1997) career trajectories and landmarks, experts decline more quickly but they also begin to stabilize much earlier and at a higher skill level than their less able colleagues. The crucial mediating role in this process is played by activity. Experts may lose more, but the accumulated knowledge obtained through practice is helping them to preserve their skill level as they get older. Age is cruel to the more able, but it may be substantially less cruel to those who are active and had been practicing in the past.

References

Allison, P. D., Long, J. S., & Krauze, T. K. (1982). Cumulative advantage and inequality in science. *American Sociological Review, 47,* 615–625. http://dx.doi.org/10.2307/2095162

Almuhtadi, K. (2011). *Age trends in chess: Is aging kinder to the initially more able?* Retrieved from http://thesis.lakeheadu.ca:8080/handle/2453/313

Baayen, R. H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. http://dx.doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3,* 12–28.

Bates, D. (2005). Fitting linear mixed models in R. *R News, 5,* 27–30.

Bedard, J. (1989). Expertise in auditing: Myth or reality? *Accounting, Organizations and Society, 14,* 113–131. http://dx.doi.org/10.1016/0361-3682(89)90037-8

Bilalić, M., Kiesel, A., Pohl, C., Erb, M., & Grodd, W. (2011). It takes two-skilled recognition of objects engages lateral areas in both hemispheres. *PLoS ONE, 6,* e16202. http://dx.doi.org/10.1371/journal.pone.0016202

Bilalić, M., & McLeod, P. (2006). How intellectual is chess?—A reply to Howard. *Journal of Biosocial Science, 38,* 419–421. http://dx.doi.org/10.1017/S0021932005001185

Bilalić, M., & McLeod, P. (2007). Participation rates and the difference in performance of women and men in chess. *Journal of Biosocial Science, 39,* 789–793. http://dx.doi.org/10.1017/S0021932007001861

Bilalić, M., & McLeod, P. (2014). Why good thoughts block better ones. *Scientific American, 310,* 74–79. http://dx.doi.org/10.1038/scientificamerican0314-74

Bilalić, M., Smallbone, K., McLeod, P., & Gobet, F. (2009). Why are (the best) women so good at chess? Participation rates and gender differences in intellectual domains. *Proceedings Biological Sciences, 276,* 1161–1165. http://dx.doi.org/10.1098/rspb.2008.1576

Blum, J. E., & Jarvik, L. F. (1974). Intellectual performance of octogenarians as a function of education and initial ability. *Human Development, 17,* 364–375. http://dx.doi.org/10.1159/000271358

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology, 44,* 108–132. http://dx.doi.org/10.1006/jmps.1999.1279

Campitelli, G., & Gobet, F. (2008). The role of practice in chess: A longitudinal study. *Learning and Individual Differences, 18,* 446–458. http://dx.doi.org/10.1016/j.lindif.2007.11.006

Campitelli, G., & Gobet, F. (2010). Herbert Simon's decision-making approach: Investigation of cognitive processes in experts. *Review of General Psychology, 14*(4), 354.

Campitelli, G., & Gobet, F. (2011). Deliberate practice necessary but not sufficient. *Current Directions in Psychological Science, 20,* 280–285. http://dx.doi.org/10.1177/0963721411421922

Chabris, C. F., & Glickman, M. E. (2006). Sex differences in intellectual performance: Analysis of a large cohort of competitive chess players. *Psychological Science, 17,* 1040–1046. http://dx.doi.org/10.1111/j.1467-9280.2006.01828.x

Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology, 33,* 1–16. http://dx.doi.org/10.1037/h0081698

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 467–476. http://dx.doi.org/10.1037/0096-1523.7.2.467

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,* 46–51. http://dx.doi.org/10.1111/j.1467-9280.1996.tb00665.x

Charness, N., Krampe, R., & Mayr, U. (1996). The role of practice and coaching in entrepreneurial skill domains: An international comparison of life-span chess skill acquisition. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sport and games* (pp. 51–58). Mahwah, NJ: Erlbaum.

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,* 1146–1152. http://dx.doi.org/10.3758/BF03206384

Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology, 19,* 151–165. http://dx.doi.org/10.1002/acp.1106

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81. http://dx.doi.org/10.1016/0010-0285(73)90004-2

Cleveland, W. S., Grosse, E., & Shyu, M. J. (1992). Local regression model. In J. M. Chambers & T. Hastie (Eds.), *Statistical models in S* (pp. 309–376). New York, NY: Chapman and Hall.

Connors, M. H., Burns, B. D., & Campitelli, G. (2011). Expertise in complex decision making: The role of search in chess 70 years after de Groot. *Cognitive Science, 35,* 1567–1579. http://dx.doi.org/10.1111/j.1551-6709.2011.01196.x

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., . . . Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin, 92,* 135–152. http://dx.doi.org/10.1093/bmb/ldp033

de Groot, A. D., Gobet, F., & Jongman, R. W. (1996). *Studies in the heuristics of the professional eye*. Assen, Netherlands: Van Gorcum & Co.

Dennis, W. (1966). Creative productivity between the ages of 20 and 80 years. *Journal of Gerontology, 21,* 1–8. http://dx.doi.org/10.1093/geronj/21.1.1

Elo, A. E. (1978). *The rating of chess players, past and present* (Vol. 3). London, UK: B. T. Batsford Ltd. Retrieved from http://www.getcited.org/pub/101876597

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine, 79,* S70–S81. http://dx.doi.org/10.1097/00001888-200410001-00022

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

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

Ericsson, K. A., & Moxley, J. H. (2012). A critique of Howard's argument for innate limits in chess performance or why we need an account based on acquired skill and deliberate practice. *Applied Cognitive Psychology, 26,* 649–653. http://dx.doi.org/10.1002/acp.2841

Gaschler, R., Progscha, J., Smallbone, K., Ram, N., & Bilalić, M. (2014). Playing off the curve—testing quantitative predictions of skill acquisition theories in development of chess performance. *Frontiers in Psychology, 5,* 923. http://dx.doi.org/10.3389/fpsyg.2014.00923

Gingras, Y., Larivière, V., Macaluso, B., & Robitaille, J.-P. (2008). The effects of aging on researchers' publication and citation patterns. *PLoS ONE, 3,* e4048. http://dx.doi.org/10.1371/journal.pone.0004048

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31,* 1–40.

Goldstein, H. (2011). *Multilevel statistical models.* London, UK: Wiley and Sons Ltd.

Gong, Y., Ericsson, K. A., & Moxley, J. H. (2015). Recall of briefly presented chess positions and its relation to chess skill. *PLoS ONE, 10,* e0118756. http://dx.doi.org/10.1371/journal.pone.0118756

Horn, J., & Masunaga, H. (2006). A merging theory of expertise and intelligence. In Ericsson KA, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 587–611). New York, NY: Cambridge University Press http://dx.doi.org/10.1017/CBO9780511816796.034

Horner, K. L., Rushton, J. P., & Vernon, P. A. (1986). Relation between aging and research productivity of academic psychologists. *Psychology and Aging, 1,* 319–324. http://dx.doi.org/10.1037/0882-7974.1.4.319

Horton, S., Baker, J., & Schorer, J. (2008). Expertise and aging: Maintaining skills through the lifespan. *European Group for Research into Elderly and Physical Activity, 5,* 89–96. http://dx.doi.org/10.1007/s11556-008-0034-5

Howard, R. W. (2005). Objective evidence of rising population ability: A detailed examination of longitudinal chess data. *Personality and Individual Differences, 38,* 347–363. http://dx.doi.org/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,* 698–703. http://dx.doi.org/10.3758/BF03193903

Howard, R. W. (2008). Linking extreme precocity and adult eminence: A study of eight prodigies at international chess. *High Ability Studies, 19,* 117–130. http://dx.doi.org/10.1080/13598130802503991

Howard, R. W. (2009). Individual differences in expertise development over decades in a complex intellectual domain. *Memory & Cognition, 37,* 194–209. http://dx.doi.org/10.3758/MC.37.2.194

Howard, R. W. (2012). Longitudinal effects of different types of practice on the development of chess expertise. *Applied Cognitive Psychology, 26,* 359–369. http://dx.doi.org/10.1002/acp.1834

Howard, R. W. (2014a). Gender differences in intellectual performance persist at the limits of individual capabilities. *Journal of Biosocial Science, 46,* 386–404. http://dx.doi.org/10.1017/S0021932013000205

Howard, R. W. (2014b). Learning curves in highly skilled chess players: A test of the generality of the power law of practice. *Acta Psychologica, 151,* 16–23. http://dx.doi.org/10.1016/j.actpsy.2014.05.013

Knapp, M. (2010). Are participation rates sufficient to explain gender differences in chess performance? *Proceedings of the Royal Society of London B: Biological Sciences.* Advance online publication.

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,* 331–359. http://dx.doi.org/10.1037/0096-3445.125.4.331

Lehman, H. C. (1953). *Age and achievement.* Princeton, NJ: Princeton University Press.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9,* 339–355. http://dx.doi.org/10.1037/0882-7974.9.3.339

Masunaga, H., & Horn, J. (2001). Expertise and age-related changes in components of intelligence. *Psychology and Aging, 16,* 293–311.

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38,* 115–142. http://dx.doi.org/10.1037/0012-1649.38.1.115

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., . . . Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56,* 128–165.

Mireles, D. E., & Charness, N. (2002). Computational explorations of the influence of structured knowledge on age-related cognitive decline. *Psychology and Aging, 17,* 245–259. http://dx.doi.org/10.1037/0882-7974.17.2.245

Moxley, J. H., & Charness, N. (2013). Meta-analysis of age and skill effects on recalling chess positions and selecting the best move. *Psychonomic Bulletin & Review, 20,* 1017–1022. http://dx.doi.org/10.3758/s13423-013-0420-5

Moxley, J. H., Ericsson, K. A., Charness, N., & Krampe, R. T. (2012). The role of intuition and deliberative thinking in experts' superior tactical decision-making. *Cognition, 124,* 72–78. http://dx.doi.org/10.1016/j.cognition.2012.03.005

Naveh-Benjamin, M., & Old, S. R. (2008). Aging and memory. In J. H. Byrne, H. Eichenbaum, R. Menzel, H. L. Roediger, & D. Sweatt (Eds.), *Learning and memory: A comprehensive reference* (pp. 787–808). Oxford, UK: Elsevier. http://dx.doi.org/10.1016/B978-012370509-9.00133-9

Ng, T. W., & Feldman, D. C. (2008). The relationship of age to ten dimensions of job performance. *Journal of Applied Psychology, 93,* 392–423. http://dx.doi.org/10.1037/0021-9010.93.2.392

Radanović, J., & Vaci, N. (2013). Analysis of reaction times using mixed-effects models. *Primenjena Psihologija, 6,* 312–332.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science, 6,* 5–42. http://dx.doi.org/10.1111/tops.12078

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Redzanowski, U., & Glück, J. (2013). Who knows who is wise? Self-ratings and peer-ratings of wisdom. *The Journals of Gerontology, Series B: Psychological Science, 68,* 391–394.

Rennig, J., Bilalić, M., Huberle, E., Karnath, H.-O., & Himmelbach, M. (2013). The temporo-parietal junction contributes to global gestalt perception-evidence from studies in chess experts. *Frontiers in Human Neuroscience, 7,* 513. http://dx.doi.org/10.3389/fnhum.2013.00513

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107,* 358–367. http://dx.doi.org/10.1037/0033-295X.107.2.358

Roring, R. W., & Charness, N. (2007). A multilevel model analysis of expertise in chess across the life span. *Psychology and Aging, 22,* 291–299. http://dx.doi.org/10.1037/0882-7974.22.2.291

Salthouse, T. A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General, 113,* 345–371. http://dx.doi.org/10.1037/0096-3445.113.3.345

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging.* Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (2010). *Major issues in cognitive ageing*. Oxford, UK: Oxford University Press.

Salthouse, T. A., & Maurer, J. J. (1996). Aging, job performance, and career development. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 353–364). New York, NY: Academic Press.

Simonton, D. K. (1977). Creative productivity, age, and stress: A biographical time-series analysis of 10 classical composers. *Journal of Personality and Social Psychology, 35,* 791–804. http://dx.doi.org/10.1037/0022-3514.35.11.791

Simonton, D. K. (1983). Creative productivity and age: A mathematical model based on a two-step cognitive process. *Developmental Review, 3,* 97–111.

Simonton, D. K. (1988). Age and outstanding achievement: What do we know after a century of research? *Psychological Bulletin, 104,* 251–267. http://dx.doi.org/10.1037/0033-2909.104.2.251

Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review, 104,* 66–89. http://dx.doi.org/10.1037/0033-295X.104.1.66

Simonton, D. K. (2010). Creative thought as blind-variation and selective-retention: Combinatorial models of exceptional creativity. *Physics of Life Reviews, 7,* 156–179. http://dx.doi.org/10.1016/j.plrev.2010.02.002

Simonton, D. K. (2015). Thomas Edison's creative career: The multilayered trajectory of trials, errors, failures, and triumphs. *Psychology of Aesthetics, Creativity, and the Arts, 9,* 2–14. http://dx.doi.org/10.1037/a0037722

Snijders, T., & Boskers, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, UK: Sage.

Staudinger, U. M., & Glück, J. (2011). Psychological wisdom research: Commonalities and differences in a growing field. *Annual Review of Psychology, 62,* 215–241. http://dx.doi.org/10.1146/annurev.psych.121208.131659

Sturman, M. C. (2003). Searching for the inverted U-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management, 29,* 609–640. http://dx.doi.org/10.1016/S0149-2063(03)00028-X

Thorndike, E. L., Bregman, E. O., Tilton, J. W., & Woodyard, E. (1928). *Adult learning*. New York, NY: MacMillan.

Vaci, N., Gula, B., & Bilalić, M. (2014). Restricting range restricts conclusions. *Frontiers in Psychology, 5,* 569. http://dx.doi.org/10.3389/fpsyg.2014.00569

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin, 122,* 231–249. http://dx.doi.org/10.1037/0033-2909.122.3.231

Waldman, D. A., & Avolio, B. J. (1993). Aging and work performance in perspective: Contextual and developmental considerations. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management*: Vol. 11 (pp. 133–162). Greenwich, CT: JAI Press.

Wilson, R. S., Hebert, L. E., Scherr, P. A., Barnes, L. L., Mendes de Leon, C. F., & Evans, D. A. (2009). Educational attainment and cognitive decline in old age. *Neurology, 72,* 460–465. http://dx.doi.org/10.1212/01.wnl.0000341782.71418.6c

Zuckerman, H. (1977). *Scientific elite*. New York, NY: Free Press.

Appendix A

Descriptive Results and Estimated Coefficients From the Models

The main goal of the data exclusion was to reduce the possibility of *compositional* and *ecological fallacy* (Dennis, 1966; Simonton, 1977; Simonton, 1988) prior to the model estimation. These fallacies represent different aggregation errors that can arise when researchers use averaged values across many individuals to derive the shape of a typical age-related curve. Averaged decline rates across several careers do not necessarily reflect the trajectories of the individuals that make up the sample, and vice versa. This problem usually occurs due to the variability in the life span of experts and/or larger number of dropouts. In other words, by including those players who stopped participating in competitions (or even died), one is estimating coefficients for data that is biased toward particular trends in age-related effects. As well as a priori "aggressive" exclusion of the possible outliers, multilevel or mixed-effect regression analysis offers a promising way to mitigate aggregation fallacies (Goldstein, 2011; Snijders & Boskers, 2012).

We also investigated whether dropout rates and rating scores vary differently by age in the two datasets. This would result in a biased skill acquisition curve, especially when extrapolating the curve above 70 years. To investigate this problem, we computed mean number of played games, as well as the mean value and standard deviation of rating scores for each age decade. Descriptive results, in the case of the German database, show that players tend to drop out, resulting in fewer players above 70 years compared with other decades, yet the absolute number is still large—almost 10,000 players. However, all players are equally active; the average number of played games is between five and six and *SD* of rating scores is declining by age (see Table A2). The descriptive results are slightly different for the FIDE database, where *SD* of rating scores are higher in the tails of the age range and the average number of played games decreases by age (see Table A2). Overall descriptive results show that skill curves can be estimated on the range from 10 to 80 years, but one should be cautious in interpreting the curve after 70 years in the case of the FIDE database (see Table A1, A2, and A3).

(Appendices continue)

Table A1

Descriptive Statistics for the Main Variables in the Datasets

	German database				
	Min	Mean	Median	SD	Max
Age	10	37	35	18.13	80
Games	0	12.04	9	10.5	125
Stale play	0	.43	0	.63	21
Rating	200	1603	1635	358	2813

	FIDE database				
	Min	Mean	Median	SD	Max
Age	10	38	37	15.16	80
Games	0	9.62	5	13.0	212
Stale play	0	.26	0	.47	19
Rating	1500	2125	2136	178.18	2851

Table A2

Descriptive Statistics for the Main Variables in the Datasets Over Each Decade of Age

	German database						
	10–19	20–29	30–39	40–49	50–59	60–69	70–80
Mean (Rating)	1,279	1,726	1,779	1,717	1,662	1,631	1,559
SD (Rating)	384	311	282	273	252	239	226
Mean (Games)	5.1	5.4	5.7	5.8	5.8	6.0	5.8
SD (Games)	2.9	3.0	3.0	2.9	3.0	3.0	3.1
Mean (Stale)	.28	.47	.48	.50	.50	.45	.52
SD (Stale)	.52	.67	.67	.67	.65	.64	.67
No. of players	42,712	37,227	36,746	32,774	24,638	19,306	9,871

	FIDE database						
	10–19	20–29	30–39	40–49	50–59	60–69	70–80
Mean (Rating)	2,007	2,148	2,167	2,146	2,119	2,091	2,048
SD (Rating)	202	165	159	162	173	180	182
Mean (Games)	5.7	3.9	3.2	2.5	2.3	2.6	2.2
SD (Games)	8.2	8.2	7.1	5.9	5.3	5.3	4.7
Mean (Stale)	.22	.28	.28	.27	.26	.26	.26
SD (Stale)	.42	.46	.48	.47	.48	.50	.54
No. of players	30,747	37,846	36,811	33,798	21,886	11,915	5,796

Table A3

Descriptive Statistics for Intersection (Identical Players) of Two Datasets for the Main Variables in the Model

	German database				
	Min	Mean	Median	SD	Max
Age	10	35.6	34	16.75	80
Games	0	6.26	7	2.84	91
Stale play	0	.33	0	.5	13
Rating	200	1,871	1,904	306	2813

	FIDE database				
	Min	Mean	Median	SD	Max
Age	10	39.9	39	14.56	80
Games	0	4.0	3	8.48	152
Stale play	0	.24	0	.37	15
Rating	1,500	2,161	2,160	176.29	2,813

(Appendices continue)

Appendix B

Goodness of Fit Comparison for Each New Term Added in Linear Mixed-Effect Model

Model	AIC	BIC	logLik	χ^2	$Pr(>|\chi^2|)$
Random intercept	30435636	30435675	−1521781		
Age	30421893	30421945	−15210943	13745	<2.2e-16
+ Age^2	30091111	30091175	−15045550	330785	<2.2e-16
+ Age^3	30018308	30018385	−15009148	72805	<2.2e-16
+ Games	30007521	30007611	−15003754	10789	<2.2e-16
*Games	29992226	29992355	−14996103	15301	<2.2e-16
+ Stale play	29971070	29971084	−14985341	1000	<2.2e-16
Random slope for age	27775566	27775720	−13887771	2216665	<2.2e-16

Appendix C

Estimated Coefficients, Standard Errors, *t*-Values and *p*-Values for Linear Mixed-Effect Model With Factor That Coded Experts and Average Players

Effect	FIDE database				German database			
	Estimate	SE	t value	Pr(>|z|)	Estimate	SE	t value	Pr(>|z|)
Intercept	1852	1.61	1141	<.00	612.5	2.45	252.1	<.00
Age	10.92	.11	102.6	<.00	78.68	.21	357.7	<.00
Age^2	−.249	3.35e-3	−79.2	<.00	−1.905	7.34e-3	−254.2	<.00
Age^3	7.22e-4	2.98e-5	28.5	<.00	.013	6.65e-5	200.9	<.00
Games	.972	.01	54.5	<.00	−.111	.069	2.0	<.05
Age:Games	−.056	2.50e-3	−21.1	<.00	.135	.012	12.8	<.00
Age^2:Games	9.17e-4	8.92e-5	8.5	<.00	−4.96e-3	4.18e-4	−12.9	<.00
Age^3:Games	−6.08e-6	8.84e-7	−4.8	<.05	4.53e-5	4.25e-6	11.8	<.00
Exp	211.3	2.35	97.2	<.00	74.66	4.58	16.9	<.00
Stale play	−.310	.028	−11	<.05	−8.221	.08	−98.03	<.00
Age:Exp	8.298	.14	47.0	<.00	35.58	.36	98.1	<.00
Age^2:Exp	−.222	3.85e-3	−42.8	<.00	−1.166	.01	−111.4	<.00
Age^3:Exp	2.32e-3	3.51e-5	48.0	<.00	.009	1.02e-4	106.6	<.00
Games: Exp	−.415	.02	−13.5	<.00	7.448	.11	62.1	<.00
Age:Games:Exp	.090	3.10e-3	22.5	<.00	−.801	.01	−47.7	<.00
Age^2:Games:Exp	−3.09e-3	1.10e-4	−19.2	<.00	.023	6.19e-4	38.5	<.00
Age^3:Games:Exp	2.93e-5	1.11e-6	15.6	<.00	−2.08e-4	6.38e-6	−32.2	<.00

Note. Reference level in the case of both datasets were average players. Age in table represents the age of players, where Age^2 and Age^3 represent quadratic and cubic term in the model, also, Exp variable codes whether players were experts or non-experts.