

Individual differences in expertise development over decades in a complex intellectual domain

ROBERT W. HOWARD

University of New South Wales, Sydney, New South Wales, Australia

Learners acquire expertise at different rates and reach different peak performance levels. Key questions arise regarding what patterns of individual differences in expertise development occur and whether innate talent affects such development. International chess is a good test domain for both issues, because it has objective performance measures, actual practice measures (number of games), longitudinal population data, and minimal gatekeeper influence. Players' expertise development typically follows either a logarithmic or a power-function curve, approaching asymptote by around 750 games. A comparison of eventual top players and other eventually well-practiced players typically reveals a performance difference at domain entry, which widens progressively with practice and then stays large and constant. The data show various correlated signs of apparent greater natural talent in eventual top players: precocity (indexed by entering the domain and gaining the grandmaster title much younger on average), faster acquisition of expertise (indexed by fewer years and games needed to gain the grandmaster title from domain entry), and a higher peak performance level after extensive actual practice. A factor analysis found evidence for an underlying natural talent factor that constrains ultimate performance level.

Learners acquire expertise at different rates and reach different ultimate performance levels. Several key questions arise (Ackerman, 2007). One is what patterns of individual differences in expertise development occur with practice in different domains. Research suggests that, in simpler domains in which skills can be quickly automated, thus reducing the load on working memory, initial individual differences eventually may tend to converge (Hunt, 2006). In complex intellectual domains in which performance depends heavily on domain-specific knowledge, attentiveness, and use of working memory (such as an air traffic control task), initial individual differences may stay relatively constant with practice or may widen (Ackerman, 2007; Ackerman & Beier, 2006; Ackerman & Cianciolo, 2000). For instance, in mathematics education, initial differences in the skill of performing calculations may widen as material becomes progressively more difficult to understand (Ackerman, 2007).

However, relatively little is known about patterns of individual differences in performance after very large amounts of practice and/or over very long periods, such as decades. Experiments with more than a few hours or days of practice typically are impractical, and researchers must often use the vocational domain (Hunt, 2006) and correlational studies. Then, practice and performance levels can be difficult to measure, and dropouts and gatekeeper influences can make results difficult to interpret.

A second, related question is whether innate talent affects expertise development and limits ultimate performance level. Views are polarized (Howe, Davidson, &

Sloboda, 1998; Hunt, 2006). The longstanding "natural talent" view acknowledges that practice is essential to instill knowledge and skills, but holds that individual differences in partially innate abilities and other traits affect learning and performance (Detterman, Gabriel, & Ruthsatz, 1998) and constrain ultimate performance levels. Francis Galton (1869/1979) argued that natural talent differences among individuals set differing, fixed asymptotic performance limits. Training can improve skill only to a certain level, and then the level of "maximum performance becomes a rigidly determinant quantity"; there is an immutable limit on performance, which the individual "cannot by any education or exertion overpass" (Galton, 1869/1979, p. 15).

The opposing "expert performance" view holds that expertise reflects only extensive practice (Ericsson, 2006; Ericsson & Charness, 1994). Top performers have no greater natural talent; they simply have practiced more, and any "normal, healthy individual" can reach elite performance levels by starting early enough and practicing extensively (Ericsson, Roring, & Nandagopal, 2007). However, individuals may need high-quality "deliberate practice," defined below, to get beyond performance asymptotes.

What Is Natural Talent?

Natural talent is difficult to define (Howe et al., 1998; Simonton, 2008). The concept derives from common everyday observations of individual differences in skill acquisition. Some people acquire a skill much more rapidly than do others. Some reach high performance levels with seemingly little effort and time. A few achieve very high

R. W. Howard, rwh@unsw.edu.au

performance levels at a much younger age than do most others. Take a hypothetical example of 10 people starting to learn touch-typing. One or two might grasp the task easily and start at a higher typing speed than the others, improve typing speed at a faster rate, and eventually reach a higher asymptotic typing speed. Perhaps one person never learns to touch-type at all, despite much practice. The rest lie somewhere in between. The highest achiever is seen as having more natural talent than the others, and the lowest achiever is seen as having less. Precocity, the acquisition of a high level of expertise at a much younger age than average, is usually seen as a sign of natural talent. Malcolm Gladwell summarized this notion as follows (although criticizing it): “We take it as an article of faith . . . that great ability . . . is invariably manifested early on, that to be precocious . . . [is] a predictor of future success” (see Wargo, 2006). Gagné (2007), arguing for natural talent for music, cited the musical precocity of Sarah Chang as a sign of great talent manifested early on. She began violin studies at age 4 and by age 8 played well enough to secure engagements with top orchestras.

Howe et al. (1998) listed five criteria for natural talent in a domain. First, the talent is partly innate. Second, its full effects are not necessarily evident early, but there are some advance indications before exceptional levels of mature performance have been demonstrated (e.g., eventual top performers should show some precocity). Third, these early indications allow prediction of who is likely to excel (e.g., the precocious should reach higher ultimate performance levels). Fourth, only a minority are talented (e.g., only some participants in a domain show such signs). Fifth, talents are relatively domain specific. Consider height in basketball. Height is partly innate, and advantages of greater height are evident early and give a very tall player a consistent advantage from domain entry. Signs such as precocity and faster learning correlate with ultimate performance level.

Simonton (1999, 2008) argues that natural talent is best seen as part of a process that involves experience, as a feature of natural endowment that has one or both of the following effects: (1) Talent enhances training by allowing a talented person to attain a higher level of expertise for a given unit of training or by allowing mastery of the skill in less training time than average. (2) Talent enhances performance. Talented individuals with a given amount of accumulated practice exhibit a higher level of output than do the less talented with the same level of accumulated practice.

Natural talent in a domain might consist of several ability and personality factors. Simonton (2008, p. 31), discussing possible natural talent for scientific research, notes, “It is extremely unlikely that endowment constitutes a homogeneous psychological capacity. . . . Instead the natural endowment most likely consists of a weighted composite of numerous and highly specific intellectual and personality characteristics.” In basketball, natural talent might include both height and speed. In intellectual domains, it might include IQ, specific abilities, creativity, and perhaps high motivation. Barron and Harrington (1981) argue that, for high achievement, complex intel-

lectual domains may require only a threshold IQ score of around 120, beyond which creativity and motivation differences create performance differences, although not all agree (e.g., Kim, 2005). Thus, IQ might correlate highly with chess skill in players from a wide IQ range (e.g., 70–130), but it would correlate little if all players were over IQ 120.

To summarize, expected signs of natural talent in a domain are precocity, faster acquisition of skill than average, which allows quick ascent to a higher performance level, and an eventual performance ceiling higher than that of the less talented who get the same amount of practice. Indexes of these signs should correlate with each other, if they derive from an underlying factor (natural talent).

What Is Practice?

Researchers have defined *practice* differently. Psychologists in the traditional skills research area have defined it as the actual performance of a task, such as batting a ball, driving a car, drawing from a mirror image, or playing chess (Downey & Anderson, 1917; Smyth, 1975). This definition is implicit in the power law of practice, which interrelates skill performance and amount of practice (Heathcote, Brown, & Mewhort, 2000; Howard, 1995; Shrager, Hogg, & Huberman, 1988). This law holds that the relationship between performance and amount of practice follows a power function: Improvement is most rapid in the early stages of training; then, performance improvements diminish progressively as performance approaches asymptote (Heathcote et al., 2000). A *power function* implies “a learning process in which some mechanism is slowing down the rate of learning” (Newell & Rosenbloom, 1981, p. 18), so that learners eventually approach asymptotic performance levels.

Here, this traditional sense is termed *actual practice*—that is, actual performance of a task. Actual practice is very important. Knowledge acquired from books and teachers must be applied and performed; no one can adeptly drive a car, perform surgery, or fly a jet without extensive actual practice.

There is also *deliberate practice* (Charness, Krampe, & Mayr, 1996; Ericsson, 2006), which Charness, Tuffiash, Krampe, Reingold, and Vasyukova (2005, p. 152) define as the performance of “appropriately challenging tasks that are chosen with the goal of improving a particular skill.” Ericsson and Ward (2007, p. 349) offer “a (typically planned) training activity aimed at reaching a level just beyond the currently attainable level of performance by engaging in full concentration, analysis after feedback, and repetitions with refinement.” In chess, deliberate practice might involve receiving coaching, reading chess books, studying opening variations and endgame positions, studying games of top players, and playing against a very strong computer to identify one’s weaknesses. Deliberate practice has been differentiated from one type of actual practice, competition (e.g., competing in the Olympics every 4 years), and from play (e.g., playing offhand chess games for fun). Only actual practice could be measured in the present study.

Evidence on the Natural Talent Issue

The issue of natural talent is important. If amount of practice is key, ability testing could be de-emphasized and views of giftedness would change (Ericsson, Nandagopal, & Roring, 2005). The expert performance view has entered the public domain as advice to parents and coaches (Ross, 2006) and seems increasingly widely held. Bilalic, McLeod, and Gobet (2007) query whether chess play even needs intelligence.

Few would dispute that practice is a powerful and important factor in expertise development. However, that practice is the only significant variable and that any "normal, healthy individual" can reach elite performance levels contradicts everyday experience, a century of use of psychometric ability tests, and studies in the skills learning tradition (Ackerman, 2007). The expert performance view requires extraordinary evidence but is difficult to test, because nature and nurture notoriously are difficult to separate and methodological problems of existing studies abound (Sternberg, 1996). As mentioned, experiments usually are impracticable over long intervals. Practice levels can be difficult to measure. Studies are often retrospective and correlational and rest on experts' subjective estimates of practice hours from the start of "serious practice" (e.g., joining a chess club, taking violin lessons). Such data are sometimes backed up with diary entries and parental reports but often are not. Memory is notoriously unreliable, and some studies have shown that such self-estimates can be quite inaccurate (Deakin, Cote, & Harvey, 2006; Hodges & Starkes, 1996). Indeed, this measurement problem is sometimes acknowledged. For instance, Charness et al. (2005, p. 162) relied on such self-estimates, but said "we cannot ascertain the degree to which different skill groups or age groups in our study may have overestimated the frequency of duration of time spent on various chess activities."

Furthermore, retrospective correlational studies may have sampling problems, lacking an appropriate control group (Sternberg, 1996). The less talented may drop out early, and identified experts with extensive practice may be atypical of all who start participating in a domain. Experts may just be those with much natural talent who persisted because they succeeded early (Detterman & Spry, 1988). As Gardner (1995, p. 803) put it, "If one plays dozens of games of chess but never wins [or] if one practices piano for several months but is unable to remember a piece from one day to the next . . . then how likely is it that one will persevere in such an endeavor?" Do some unidentified individuals get much practice but never become experts, or do only the very talented persist? Few domains are truly meritocratic, and gatekeeper influences may partly mask natural talent differences. The very talented can be out-competed by the very well connected, who also get greater time and resources for extensive practice.

Space limitations preclude an extensive review of relevant research (see Ericsson, 2006; Ericsson et al., 2007). High practice levels can increase skill level greatly. Top performers often report starting at a younger age and having had more practice than that reported by poorer performers (Charness et al., 1996; Ericsson, 2006). An

uncontrolled illustration of the evident power of practice can be found in the three Polgar sisters, whose father selected chess as a domain to test nature/nurture views. All three were home-schooled and were exposed to much chess study from an early age, reportedly studying chess for 5–10 hours a day (Charness et al., 1996). All became elite performers, two (Judit & Zsuzsa) becoming grandmasters and one (Judit) reaching the top 10 list. But the sisters firmly believe in natural talent, stating that they thought Sofia had the greatest natural talent for chess of the three. As Zsuzsa put it in a 2005 *Psychology Today* interview, "Everything came easiest to her [Sofia] . . . but she was lazy."

Natural talent, indexed by IQ score, predicts performance in school and in jobs initially (Hunt, 1995; Ree & Earles, 1991; Schmidt & Hunter, 1998). High performers in professions typically have high IQ scores (Barron & Harrington, 1981).

Chess As a Useful Test Domain for the Natural Talent Issue

Chess is a useful domain for testing the natural talent issue, because playing chess seems to be an obviously intellectual task and because objective performance rating systems allow easy identification of chess experts (Ross, 2006). Furthermore, organized chess has relatively little gatekeeper influence and no obvious glass ceiling. Of course no intellectual domain is a complete meritocracy, because advantage to those born into higher socioeconomic classes starts early. For instance, wealthier individuals might have better access to good chess coaching, chess materials for study, and travel to and time for tournaments. However, anyone can rise quickly in the domain by playing in open tournaments and winning games. Players do not need the right social background, university degrees, connections, or favor of supervisors. There is no need for incessant networking and tireless self-promotion. Natural talent for chess might include an IQ score of 120 to reach top levels, good visuospatial ability (Howard, 2005), creativity, and perhaps extreme competitiveness. Chess expertise does require extensive domain-specific knowledge (Chase & Simon, 1973) and a lot of study.

Several studies have examined psychometric ability/chess performance links, with mixed results, but usually with small samples and often with few or no master-level players. Djakow, Petrowski, and Rudik (1927) found that eight masters at the Moscow 1925 tournament had intelligence test scores similar to those of a control group. However, one chess player held a doctorate in mathematics, and the control group was probably quite well educated (average IQ score over 120?). It would hardly have been a random sample of Russians, who at that time were predominantly peasants and factory workers. Some studies suggest a link between chess skill and IQ in children (Frydman & Lynn, 1992; Horgan & Morgan, 1990) and with some psychometric test scores in adults (Doll & Mayr, 1987; Grabner, Neubauer, & Stern, 2006; Grabner, Stern, & Neubauer, 2007). Some evidence (based on $n = 1$) suggests that IQ score is not particularly important beyond 120: The news magazine *Der Spiegel* (1987)

found that then world champion Garry Kasparov, widely considered to be the greatest player ever, scored not much above 120. Waters, Gobet, and Leyden (2002) found little link between chess skill and scores on the shape memory test, which measures imagery vividness. However, ability to manipulate mental images, not assessed by this test, may be more important, as these authors did note.

The Present Approach and Data Set

More research is needed on the question of talent versus practice. The present study used a quite different approach from that of previous studies and a data set that overcomes most of those studies' difficulties. The study used population-level data from a recently available longitudinal data set of all international chess players (Howard, 2006). These data include objective performance ratings from a player's entry into the international domain, as well as an objective measure of actual practice: number of rated games. As mentioned, deliberate practice could not be measured with this data set. The game counts do not include offhand and unrated games, but these too are essentially unmeasurable. Finally, there were no sampling problems, and performance could be gauged over decades, allowing a direct test, for example, of whether some unidentified participants get much actual practice but never reach top levels. It was possible to see what might be going on with all participants in an important domain with little gatekeeper influence.

The natural talent view makes many clear, testable predictions about how natural talent for chess should manifest itself. Howe et al.'s (1998) criteria and Galton's (1869/1979) notion of differing, immutable performance limits suggest key signs, and the present study looked for them. Specifically, do eventual top players show precocity (indexed by entering the international domain and gaining the grandmaster title much younger on average) and faster acquisition of expertise than average (measured by years and number of games needed to gain the grandmaster title)? If such signs of precocity and speed of expertise acquisition derive from an underlying natural talent factor, they should correlate with each other and with peak performance level (measured by peak rating). The present study also determined whether players do reach widely differing asymptotic performance levels with equivalent amounts of actual practice, this perhaps being evidence of natural talent's setting Galton's "immutable performance limits."

Aims of the Present Study

The present study had two aims. The primary aim was to see what patterns of expertise development occur over decades and with extensive actual practice in a complex intellectual domain. Very little is known about what might happen over such long intervals. Specifically, does the power law of practice apply over very long intervals and with extensive amounts of practice? Do expertise development curves fit power functions, as other studies have found? What patterns of individual differences occur? Does performance of all players converge after extensive practice, as often occurs in the vocational domain? Or,

because chess is a complex intellectual domain with continual demands on attention and working memory (like mathematics or Ackerman's air traffic control task), would initial differences stay constant or even widen?

The second, related aim was to see whether the predictions of the natural talent view would hold up. The natural talent view makes very clear, testable predictions about what ought to happen. If few of these predictions should be found to actually hold up in chess, it would be evidence against the existence of natural talent in chess, or at least, it would be evidence that extensive practice might indeed be the key factor.

INTERNATIONAL CHESS FEDERATION RATING DATA AND THE CHESS ENVIRONMENT

The International Chess Federation (the *Fédération Internationale des Échecs*, or FIDE), which administers international chess, designates certain strong tournaments or matches as international ones and rates those tournaments' games. Tournaments often have between 7 and 11 rounds. Often 1 or 2 rounds are played per day, and a game might last 4 h or more.

International chess dates back centuries, but FIDE has rated games only from around 1970, when the first FIDE rating list appeared. This list covered games in international events in the previous 2 years or so. FIDE initially issued one list a year, then two (in January and July); in 2000, it issued three, and since then, four. The rating lists include all players active in international tournaments since 1970 who have achieved a minimum performance level. Players in rated tournaments who lack an initial FIDE rating get on the FIDE list if they achieve a specified minimum performance level but later may be dropped from the list if they become inactive or fall below the minimum rating level. The lists give any title gained (e.g., grandmaster), as well as a player's current performance rating on a scale from about 1400 to about 3000 (Elo, 1986). A player's rating may change with each game according to the game's result and the opponent's rating, so that the rating will eventually measure a player's skill at a given time. For example, if a player rated 2300 beats one rated 2400, his rating rises by some points, whereas his opponent's rating falls by the same number. Defeating a much higher rated player adds more, up to a maximum of 25 points for a single game. Defeating a much lower rated player adds relatively few points. A series of wins against higher rated players raises a player's rating progressively, whereas a series of losses to lower rated players decreases it. A rating is considered reliable after around 25 initial games (Elo, 1986). The number of games played in a rating period has been tallied since July 1985. Howard (2006) and Elo give greater detail about ratings data and how ratings are calculated. Like almost all real-world data, the FIDE ratings data are a bit messy, and all values below should be seen as approximations (see Howard, 2006, for details).

The rules for getting on and staying on the rating list have changed over time. Males once needed a minimum rating of 2200, but FIDE dropped this requirement starting

in 1993. A 2200 rating shows a very strong player and is considered master level outside the international domain, and so players first on the list beginning in 1993 are selected much less. The mean rating in July 1989 was 2283.94 ($SD = 101.29$) and that list had 7,192 players. At that time, a 2200 rating put a player in the top 10,000 or so players out of the world's many millions. The highest ever rating is 2851, achieved by Garry Kasparov in 1999–2000.

Two of the elite categories of international chess players are the world championship candidates and grandmasters. *Candidate* derives from candidates' elimination matches with varying numbers of players, after which the winner challenges the world champion for the world title. Here, a candidate is defined as anyone who has ever been rated one of the top 10 players. The prestigious grandmaster title is awarded by FIDE after a player has met specific title requirements, which have altered slightly over the years. Nowadays, most players qualify by achieving several minimum scores (*title norms*) in international tournaments and a rating over 2500. The July 1989 list had 282 grandmasters, and the July 2006 list had 1,009. FIDE also awards several lesser open titles, with easier requirements. These include international master and the least prestigious of open titles, FIDE master.

The raw data used in the present study can be downloaded (see Howard, 2006).

OVERVIEW OF STUDIES 1–5

Because numbers of games are tallied only from July 1985, only players first on the list from that date were used in all but Study 5. Three main groups were examined. The first group consisted of the 69,998 players first on the list from July 1985 through July 2006. The second group comprised the 3,471 players first on the list from July 1985 through July 1989. All males in this group entered under the stringent 2200 minimum rating requirement, and they have a long forward career span. Ideally, only players first on the list in 1 year (e.g., 1986) would be studied, in order to give a long forward span of 20 years and to reduce the chance of cohort effects, which can occur when players first on the list much later are included. However, the sample size would then be too small. The July 1989 cutoff was chosen because, at 4.5 years, the span of inclusion is still relatively small, minimizing possible cohort effects, but the sample size was reasonable. The third group, examined in Study 2, comprised all grandmasters first on the list from July 1986 through July 1995.

Three main subgroups were used: eventual candidates (all candidates also are grandmasters), eventual noncandidate grandmasters, and players who had played at least 740 games but who had not become grandmasters by July 2006. The 740-game value was the 90th percentile (see Study 2), the count by which 90% of grandmasters had achieved the grandmaster title (and 740 games was approximately the number at which players neared an asymptotic performance level; see Study 4). A count of 740 games suggests a very active, well-practiced player. For example, Garry Kasparov, ranked first in the world from 1985 to

2006, played 761 FIDE-rated games from July 1985 to July 2006. These highly active players form a kind of control group for the grandmasters. They undertake about the same amount of actual practice and spend as much time in the domain, but possibly have less natural talent.

Study 1 measured attrition rates over number of games of players first on the list from July 1985 to July 1989 and looked for sizeable numbers of players who got extensive actual practice but did not become grandmasters. Study 2 determined the distribution of games played to get the grandmaster title and looked for such signs of natural talent in grandmasters as precocity and faster expertise acquisition as well as at whether these signs correlated with ultimate performance level. Studies 3 and 4 looked at individual differences in expertise-development patterns over time and over number of games in candidates, noncandidate grandmasters, and nongrandmasters who had played at least 740 games. Study 5 looked at the effects of extreme amounts of actual practice, over 2,300 games, to determine whether players continually improve with extensive practice.

STUDY 1

Career Practice Levels and Attrition Rates

The typical study supporting the expert performance view compares the reported amounts of practice of experts and nonexperts and finds that the experts report having had much more practice. The typical conclusion is that amount of practice is the key factor in determining ultimate performance level and that “any normal, healthy individual” can reach elite performance levels (i.e., become an expert) by starting early enough and practicing extensively. Ericsson et al. (2007) have reviewed many such studies. One example of such a study is Charness et al. (1996); they found that more skilled chess players reported more practice than did less skilled ones and concluded that “as the literature on skill in music performance revealed, high-level performance is a function of a very high investment in deliberate practice alone” (p. 77).

However, as mentioned, it has been argued that such studies lack a control condition (Sternberg, 1996)—say, of individuals who do not become experts, but who get at least as much practice as experts do. The expert performance view might counter that, because extent of practice is the only key factor, few or no such persons exist. Anyone who gets extensive practice will become an expert. This argument can be evaluated by determining whether many individuals stay in a domain for decades and get extensive actual practice, yet never achieve expert status. Finding many such individuals, say at least 50% of all participants, would be evidence against the expert performance view.

However, results of such a test might be hard to interpret. If 50% of participants who persist and get extensive practice do not become experts, this would be evidence against the expert performance view. But if few or no such individuals are found, there are at least two interpretations. As mentioned, natural talent, motivation, and amount of

practice may all be confounded (Sternberg, 1996). Only the talented and initially successful may stay in the domain and get extensive practice.

The FIDE data include all rated participants in the international domain and allow such a test. The group first on the list from July 1985 to July 1989 was used, because a long forward time span was needed. However, most players entering the list were rated at least 2200 and were already experts. So a very stringent criterion of achieving “expert” status was used: gaining the grandmaster title.

Study 1 had two aims. The first was to determine whether all players who persist in international chess and get extensive actual practice become grandmasters, or whether many individuals stay in the domain, get extensive practice, yet never become grandmasters. The second aim was to gauge attrition rates over number of games, to determine whether high dropout rates over number of games would be an important confounding factor in later studies.

Method

Participants. Participants comprised all 3,471 players who first entered the list from July 1985 through July 1989. They had many nationalities, and about 9.9% were female. The players were divided into three groups: eventual candidates ($n = 8$, 1 female), eventual noncandidate grandmasters ($n = 168$, 3 female), and nongrandmasters by July 2006 ($n = 3,297$).

Procedure. All players who became grandmasters by July 2006 were identified from “g” in the title field (Howard, 2006). Date of obtaining the title was assigned as that of the first list in which “g” appeared. This procedure may overestimate the actual date of meeting requirements. For female grandmasters, FIDE did not put “g” in the title field in most lists; their title date was determined from media reports and was assigned as the next rating list date. Data of one grandmaster who achieved the title under allegedly suspect conditions were not used.

Total games played from list entry to July 2006 was computed for each player. Game totals for each player were placed into 50-game categories (e.g., 0–49, 50–99, etc.), and the percentage of players in each group who had reached each successive game total category was computed. Very few played more than 1,450 games, and the final category used was 1,400–1,449 games, which had only 33 players. Very active players might skip a game category, because FIDE reports ratings and game counts every few months. Attrition

rate over number of games was defined as the percentage of players who remained in each successive game category.

Results

The median game total for all players was only 109. The median was 1,125.5 for candidates, 1,000.5 for grandmasters, 997 for noncandidate grandmasters, and 942 for non-grandmasters who had played at least 740 games.

Figure 1 presents attrition rates over number of games for candidates, noncandidate grandmasters, all nongrandmasters, and all players. Figure 1 shows that rates of attrition over number of games are low for candidates and noncandidate grandmasters and high for nongrandmasters. By the 900–949 game category, only about 1.5% of nongrandmasters were left, compared with more than 55% of candidates and noncandidate grandmasters.

However, attrition over time was less severe. Most players were still active in the domain more than a decade after entry. A total of 2,182 (62.86%) of all players played at least one game from January 2001 to July 2006, and 2,010 (60.96%) of the nongrandmasters did.

The median number of games for becoming a grandmaster for all grandmasters first on the list from July 1985 to July 2006 was 390, and the 90th percentile was about 740 (see Study 2). Only about 11.59% of nongrandmasters first on the list from July 1985 to July 1989 played 390 games or more. Of the 220 players who played at least 740 games, 131 (59.55%) became grandmasters. Of the remaining 89 players, 74 (about 83%) achieved the lesser, but still prestigious, international master title by July 2006.

Discussion

The prediction of the expert performance view was partly supported. Most players did not get enough actual practice to have a realistic chance of becoming a grandmaster, and most players who did get extensive actual practice (defined as over 740 games) did become grandmasters. However, a substantial percentage (about 40%) did not, and these players formed a control group in Study 3. Furthermore, the finding has alternative interpretations, as mentioned above.

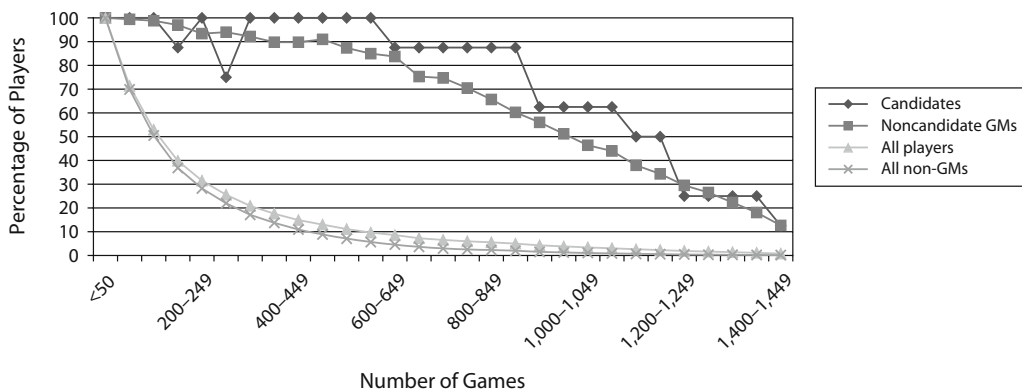


Figure 1. Attrition rates over number of games in successive 50-game categories of players first on the list from July 1985 through July 1989. GM, grandmaster.

The high attrition rate over number of games for non-grandmasters is surprising, because most players are high performers. As mentioned, males needed a minimum rating of 2200 to get on the list. This high attrition over number of games makes interpretation of findings about expertise development difficult, because so few get enough actual practice to reach asymptotic performance levels. Those who do may be very unrepresentative of all participants.

STUDY 2 Signs of Greater Natural Talent in Grandmasters and Candidates?

Study 2 looked for signs of natural talent in grandmasters. As mentioned, expected signs of natural talent are precocity, faster acquisition of expertise than average, and a higher eventual performance ceiling than those with less talent who practice as much. Those with greater natural talent should perform better at earlier ages and should consistently outperform those with less talent but with similar amounts of actual practice. The more talented should reach a higher performance level with the same amount of actual practice and should have a higher peak performance level. Furthermore, if these signs represent effects of an underlying characteristic (natural talent), they should correlate with each other.

Study 2 measured the signs as follows: *Precocity* was measured by age of entering the list and of gaining the grandmaster title. *Speed of expertise* acquisition was measured by years and number of games needed to gain the grandmaster title from domain entry. *Peak performance level* was measured by peak rating. A factor analysis was conducted to determine whether scores on the measures correlate with each other and load on an underlying factor.

Furthermore, because natural talent has a genetic component, it might be expected to have a roughly normal distribution, as do many human characteristics, such as blood pressure, height, and IQ (Gregory, 2004; Talwar & Srivastava, 2004), even at the top of the skill distribution. As McIntire and Miller (2000) put it, “Most distributions of human traits—from height and weight to aptitudes and

personality characteristics—form a normal curve” (p. 85). Scores on a measure of natural talent, such as number of games to the grandmaster title (an index of speed of expertise acquisition), might then be expected to be roughly distributed normally. (Study 2 actually found that games to title is a good measure of the underlying factor, as is peak rating.) If only amount of deliberate practice is key, there is no obvious reason to predict a normal distribution of games to the title, though it might be possible to think of one. However, the natural talent view makes a clear, testable prediction.

Study 2 tested these predictions. Do eventual top players enter the list younger and gain the grandmaster title faster, younger, and with fewer games? Do the latter indexes of precocity and speed of expertise acquisition correlate with peak rating? Is number of games to the grandmaster title normally distributed?

Method

Participants. This study considered all 69,998 players (about 7.28% female) first on the list from July 1985 to July 2006. One subgroup comprised all 591 players who had gained the grandmaster title by July 2006 (8 were female). Another subgroup comprised the 435 grandmasters on the list from July 1985 to July 1995 with known birth dates (5 were female).

Procedure. Number of games to attain the grandmaster title for each player was calculated from totals in the games field from entry onto the list up through the list in which a “g” first appeared. Years to gain the title was calculated similarly. Other statistics were computed as described below.

Results

Figure 2 presents the distribution of games to attain the title for all grandmasters first on the list from July 1985 to July 2006 in 100 game categories. The distribution appears somewhat skewed but is only just outside an excellent approximation to normality, which is usually considered to be within boundaries of 1 and -1 for skewness and kurtosis (skewness = 1.11, kurtosis = 1.37). Values between 2 and -2 are often considered reasonable approximations to normality. The median is 390 games, and the 90th percentile is around 740 ($M = 435.2$ games, $SD = 215.4$).

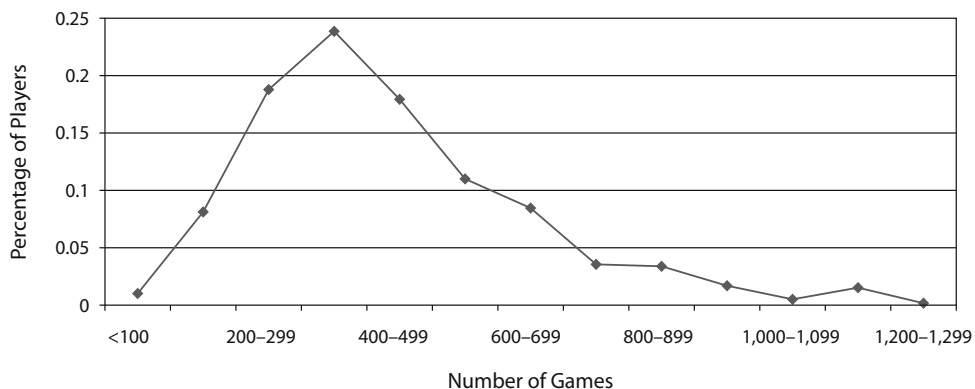


Figure 2. Distribution of games needed to attain the grandmaster title for all grandmasters first on the list from July 1985 through July 2006.

Table 1
Data of All Players Who Entered the List Between July 1985 and July 2006

	<i>n</i>	Age First on List		Mean Peak Rating	Mean Peak Rank	Med. Total Games	Games to Title		Years to Title		Age at Title	
		Med.	<i>SD</i>				Med.	<i>SD</i>	Med.	<i>SD</i>	Med.	<i>SD</i>
Candidates	15	13.58	2.45	2756.13	4.07	866	212	73.97	3.50	1.14	17.33	1.92
All GMs	591	16.33	5.25	2575.59	229.98	722	390	215.32	7.50	3.57	24.00	6.68
Non-GMs (740+)	181	19.62	6.81	2451.38	1,052.87	921						
All Players	69,998	22.75	15.15	2135.78	27,128.00	30						

Note—Birth dates of all candidates were available, but birth dates were available for only about 96% of grandmasters (GMs), 96% of non-GMs who had played at least 740 games, and about 76% of all players.

The criterion of 740 games was used in Studies 3 and 4 to select players with extensive actual practice but no grandmaster title by July 2006.

Table 1 presents career data for all players, and shows clear evidence of precocity in eventual top players. Grandmasters are the youngest to get on the list [$F(1,500) = 474.564, p < .001$]. Grandmasters get on the list younger on average than nongrandmasters [$F(1,567) = 8.006, p < .01$], get the title in fewer years [$F(1,589) = 24.519, p < .001$], with fewer games [$F(1,589) = 15.044, p < .001$], and at an earlier age [$F(1,567) = 23.53, p < .001$].

The natural talent view would predict that a natural talent factor, which limits ultimate performance level, lies behind this evidence of precocity and faster rate of expertise acquisition in eventual top players. A factor analysis was conducted to test for such evidence. Table 2 presents a correlation matrix for games to title, age at title, and peak rating for the 435 grandmasters with known birth dates first on the list from July 1985 through July 1995. Using July 1995 gives a good sample size and a long forward span of at least 11 years for players to approach peak rating. Including age on list and/or years to title in the analysis yields collinearity (between age on list and age at title, and between games to title and years to title) and

a Kaiser–Meyer–Olkin measure below 0.5, and thus the data would be unsuitable for a factor analysis. Age on list may be more of an accidental factor than age at title, so the latter was retained. Similarly, games to title was preferred, because years to title may include long breaks from chess. Table 2 shows significant correlations among peak rating, games to title, and age at title. Table 3 presents results of a factor analysis (principal components extraction). The factor analysis yielded a single factor with an eigenvalue over 1, which accounts for about 63% of the variance. All three variables load highly on this factor.

Discussion

Predictions of the natural talent view were met. The distribution of games to title was very roughly normal. Measures of rate of expertise acquisition, precocity, and peak performance levels all correlated, as the natural talent view would predict. The factor analysis suggested a common factor behind the correlations, again as the view predicted. Evidence of the factor emerged early in eventual top players (meeting one of Howe et al.'s, 1998, criteria), and peak rating loaded highly on this factor. However, this factor could be interpreted as amount of deliberate practice. This possibility seems unlikely, because evidence of it emerges early, but deliberate practice could not be measured, and that interpretation cannot be tested here.

Table 2
Correlations Between Variables for All Grandmasters With Known Birth Dates Entering the List From July 1985 to July 1995 (N = 435)

	Peak Rating	Games to Title	Age at Title
Peak rating	–	–.417*	–.549*
Games to title			.380*

* $p < .001$.

Table 3
Factor Analysis Results for Data of All Grandmasters With Known Birth Dates First on the List From July 1985 to July 1995 (N = 435)

Factor	Eigenvalue	% Variance Explained
1	1.901	63.371
2	0.65	21.668
Factor 1 Loadings		
Peak rating	–0.836	
Games to title	0.731	
Age at title	0.817	

Note—Determinant = 0.555. Kaiser–Meyer–Olkin measure = 0.654. Bartlett test of sphericity: $\chi^2(3) = 254.71, p < .001$.

STUDY 3
Individual Differences Over Time

Study 3 looked at overall patterns of expertise development over time in candidates, noncandidate grandmasters, and well-practiced nongrandmasters. The first aim was to determine what patterns of individual differences emerge over a long period and whether power functions would occur. The second aim was to look for such signs of natural talent in eventual top players as early, consistent performance superiority over time and whether players hit widely differing asymptotic performance levels. If native talent has an impact throughout expertise development in this complex domain, performance in the groups should converge little with practice. Candidates should consistently outperform other grandmasters, and grandmasters should consistently outperform nongrandmasters.

Only players first on the list from July 1985 to July 1989 were used, because of the need for a long forward time span.

Method

Participants. From players first on the list from July 1985 to July 1989, all 8 candidates (1 female), all 168 noncandidate grandmasters (3 female), and the 89 nongrandmasters (14 female) who had played at least 740 games were selected. All grandmasters were selected, regardless of how many games they had played (some had played fewer than 740).

Procedure. Because players entered the list at different times (e.g., some in July 1985, some in January 1987, etc.), ratings for each player at 6-month intervals from initial entry point were used. Thus, the first list on which a player appeared was designated Point 1, the second was Point 2, and so on. Data were used only up until the point at which all players had a data point, Point 35, covering about 17 years. In addition, the number of lists increased to three in 2000 and to four thereafter. To keep intervals between points equal, only January and July lists were used, with one exception. The July 2002 list had some missing ratings and was replaced by the April 2002 list, which had none. Overall, about 0.5% of ratings were missing (it is probable that this is mainly because FIDE dropped inactive players from the list), and these values were estimated for each player by taking the mean of values in the player's adjacent January or July lists.

Mean ratings at each point were computed for the three subgroups and for these subgroups combined (designated as *all players*).

Results

Figure 3 presents mean ratings at each point for the three groups. It suggests power functions and initial differences between groups, differences that progressively widen and then stay large and constant. Candidates and noncandidate grandmasters show an initial difference of about 0.7 *SDs*, which widens to about 2 *SDs*, suggesting that candidates acquired expertise at a faster rate. Noncandidate grandmasters and nongrandmasters show an initial difference of about 0.6 *SDs*, which widens to about 1.3. All groups apparently approached a near asymptote by about Point 25 (about 12 years from list entry), and differences thereafter stayed large and constant.

A statistical analysis confirms these impressions. A two-factor ANOVA on group (candidate, noncandidate grandmaster, and nongrandmaster) and point on list showed a highly significant effect of group [$F(2,260) = 209.807, p < .001$] and of point, with Greenhouse–Geisser correction [$F(5.529,1437.463) = 159.838, p < .001$]. The inter-

action between group and point with Greenhouse–Geisser correction was also highly significant, confirming varying differences at successive levels [$F(11.057,1437.463) = 11.611, p < .001$].

The same overall pattern occurred when candidates were compared with noncandidate grandmasters, when all grandmasters (candidates and noncandidate grandmasters) were compared with nongrandmasters, and when noncandidate grandmasters were compared with nongrandmasters. In a comparison of candidates with noncandidate grandmasters, the group difference was highly significant [$F(1,172) = 108.558, p < .001$], and, with Greenhouse–Geisser correction, so were the effect of point [$F(4.589,789.277) = 111.847, p < .001$] and the interaction [$F(4.589,789.277) = 7.99, p < .001$]. In a comparison of all grandmasters (candidates and noncandidate grandmasters) with nongrandmasters, the group difference was highly significant [$F(1,261) = 227.906, p < .001$], and, with Greenhouse–Geisser correction, so were the effect of point [$F(5.362,1399.368) = 351.582, p < .001$] and the interaction [$F(5.362,1399.368) = 14.984, p < .001$]. In a comparison of noncandidate grandmasters with nongrandmasters, the group difference was highly significant [$F(1,253) = 265.135, p < .001$], and, with Greenhouse–Geisser correction, so were the effect of point [$F(5.614,1420.432) = 344.692, p < .001$] and the interaction [$F(5.614,1420.432) = 12.915, p < .001$].

The significant interaction shown in Figure 3 suggests that the effect of group is different at different points. Are there no significant group differences at some points (e.g., at the start), or do group differences just vary in magnitude, as Figure 3 suggests? Because there were so many points, only the starting point and/or the point of smallest difference between pairs of groups were examined in post hoc tests. The smallest difference between noncandidate grandmasters and nongrandmasters is at Point 1, but the smallest difference between candidates and noncandidate grandmasters is at Point 3. Group differences are present at the start and at all levels. An overall ANOVA with ratings of the three groups at Point 1 showed a highly

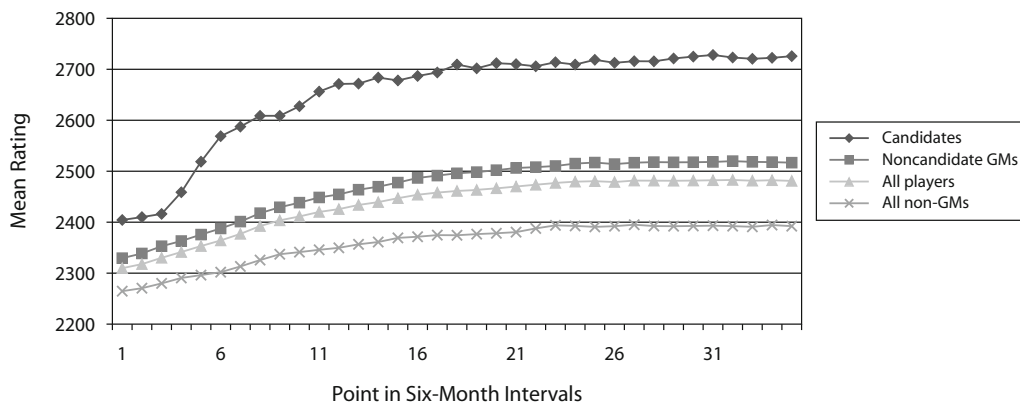


Figure 3. Rating trajectories over time from list entry (Point 1) of all players in various categories first on the list from July 1985 through July 1989. GM, grandmaster.

significant effect of group [$F(2,260) = 26.052, p < .001$]. A post hoc Tukey HSD test showed that ratings of all pairs of groups at Point 1 were significantly different: candidates and noncandidate grandmasters ($SE = 28.226, p < .05$), noncandidate grandmasters and nongrandmasters ($SE = 10.245, p < .001$), and candidates and nongrandmasters ($SE = 28.782, p < .001$). The smallest rating difference between candidates and noncandidate grandmasters occurs at Point 3. An overall ANOVA on Point 3 data for the three groups was significant [$F(2,260) = 34.651, p < .001$] and a Tukey HSD test showed that all pairs of groups differed significantly: candidates and noncandidate grandmasters ($SE = 26.436, p < .05$), noncandidate grandmasters and nongrandmasters ($SE = 9.595, p < .001$), and candidates and nongrandmasters ($SE = 26.956, p < .001$).

Were the three groups at asymptote with constant differences at Point 25 and beyond? An ANOVA using ratings from Point 25 onward for the three groups showed that differences really were constant and ratings were at asymptote. The ANOVA showed a highly significant effect of group [$F(2,260) = 242.179, p < .001$], and, with Greenhouse–Geisser correction, showed no significant effect of point [$F(3.92,1019.22) = 0.418$] and no significant interaction [$F(7.84,1019.19) = 0.312$].

Therefore, the interaction illustrated in Figure 3 data may be characterized as follows. Rating differences between all pairs of groups were present at all points but were smallest near the start, then progressively widened, and, by Point 25, each group reached asymptote and rating differences thereafter stayed constant.

A least squares analysis was performed on the four group curves in Figure 3 to test fits to power, linear, logarithmic, growth, and exponential models. The power law of practice would predict a power function as best fit. The curves show an excellent fit to a power or logarithmic function model, suggesting approach to an asymptote. For all players, the best fit was to a power function (adjusted $R^2 = .959$) but only marginally better than for a logarithmic (adjusted $R^2 = .957$). For candidates, the best fit was logarithmic (adjusted $R^2 = .932$) but only marginally (adjusted R^2 for power = .929). For noncandidate grandmasters, the best fit was power (adjusted $R^2 = .957$) but only marginally better than logarithmic (adjusted $R^2 = .956$). For nongrandmasters, the best fit was power (adjusted $R^2 = .955$) but was only marginally better than logarithmic (adjusted $R^2 = .954$).

Discussion

The patterns are clear cut. The curves follow power functions mostly, as predicted by the power law of practice. Initial group differences widened and eventually stayed large and constant and at asymptote. The data suggest that effects of an apparent natural talent factor emerge very early, as predicted. The candidates entered the list at a median age of 13.87 years and at a slightly higher rating level than that which the nongrandmasters achieved after nearly two decades in the domain. Players also apparently strike performance limits, which one “cannot by any education or exertion overpass.” The expert performance view might argue that the noncandidates did not

undertake the difficult deliberate practice needed to get beyond such asymptotes, but there is no evidence for such complacency.

STUDY 4

Individual Differences Over Number of Games

Study 4 looked at individual differences over amount of actual practice as measured by number of games. All players on the list from July 1985 to July 2006 who had played at least 850 games were used. Few had played more than 900 games and sample size would have been too small if additional games categories were used.

Method

Participants. Participants were all of the players first on the list between July 1985 and July 2006 who had played at least 850 games. There were 9 candidates (1 female), 201 noncandidate grandmasters (2 female), and 117 nongrandmasters (22 female).

Procedure. For each player, game totals from July 1985 to July 2006 were divided into 50-game categories (e.g., 0–49, 50–99, etc.) and the final rating of each player in each 50-game category was determined. For instance, say a player had the following data in successive lists: 2321 rating and 123 cumulative games, and 2360 rating and 149 games. Only 2360 was used. This method was used because FIDE rating lists appear only every few months and do not give a rating after each game. A list gives a final rating only after completion of all games in the previous rating period. About 0.6% of values were missing and were estimated by taking the average of adjacent cells for each player. Mean ratings in each group and subgroup were computed for categories up to the 850–899 games category.

There were no dropouts. All players in each subgroup had a value in the 850–899 games category.

Results

Figure 4 presents data for each group and for the three groups combined (all players). The patterns are similar to those for time in Study 3. The effect of group was highly significant [$F(2,324) = 269.193, p < .001$], and, with Greenhouse–Geisser correction, so were that of games category [$F(5.168,1674.455) = 98.831, p < .001$] and the interaction [$F(10.336,1674.455) = 6.627, p < .001$].

As in Study 3, the same overall patterns of group differences and an interaction occurred in comparisons of pairs of subgroups. In a comparison of candidates with noncandidate grandmasters, the group difference was highly significant [$F(1,208) = 79.265, p < .001$], and, with Greenhouse–Geisser correction, so were the effect of games category [$F(4.083,849.212) = 57.915, p < .001$] and the interaction [$F(4.083,849.212) = 5.614, p < .001$]. In a comparison of all grandmasters (candidates and noncandidate grandmasters) with nongrandmasters, the group difference was highly significant [$F(1,325) = 365.218, p < .001$], and, with Greenhouse–Geisser correction, so were the effect of games category [$F(5.086,1652.829) = 218.942, p < .001$] and the interaction [$F(5.086,1652.829) = 6.634, p < .001$]. In a comparison of noncandidate grandmasters with nongrandmasters, the group difference was highly significant [$F(1,316) = 399.984, p < .001$] and, with Greenhouse–Geisser correction, so were the effect of games category

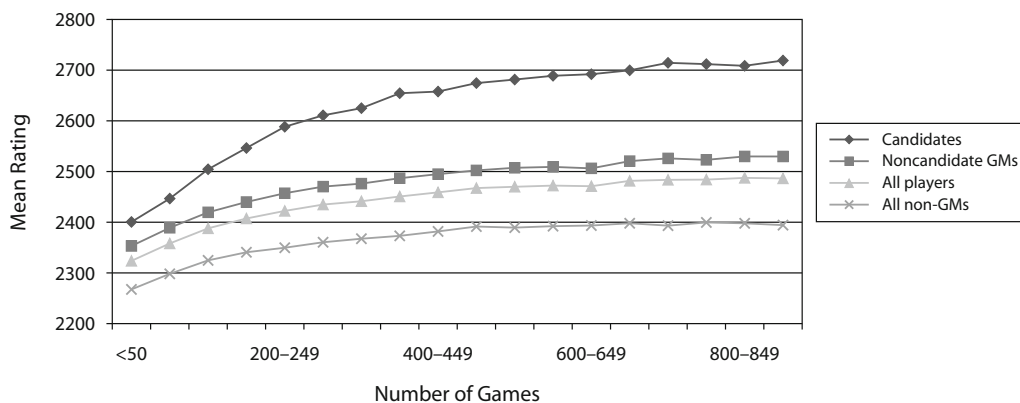


Figure 4. Rating trajectories over cumulative number of games of all players in various categories who had played at least 850 games and were first on the list from July 1985 through July 2006. GM, grandmaster.

$[F(5.131,1621.377) = 207.524, p < .001]$ and the interaction $[F(5.131,1621.377) = 5.154, p < .001]$.

The interaction found in the data in Figure 4 was explored further with post hoc Tukey HSD tests. The overall patterns were like those for time in Study 3, except that significant differences between candidates and noncandidate grandmasters emerged later, at the 100–149 games category. The smallest differences between candidates and noncandidate grandmasters and between the latter and nongrandmasters occurred at the start, in the 0–49 games category. The effect of group in the 0–49 category was highly significant $[F(2,324) = 46.242, p < .001]$. However, Tukey HSD tests showed that the difference between candidates and noncandidate grandmasters was not significant ($SD = 27.377, p < .2$), but that the difference between noncandidate grandmasters and nongrandmasters was significant ($SD = 9.344, p < .001$); and so was the difference between candidates and nongrandmasters ($SD = 27.795, p < .001$). The next smallest rating difference between candidates and noncandidate grandmasters occurred at the 50–99 games category. An overall ANOVA for all three groups at ratings in the 50–99 games category was highly significant $[F(2,324) = 56.173, p < .001]$. Candidates and noncandidate grandmasters did not differ significantly ($SE = 26.564, p < .1$), whereas noncandidate grandmasters and nongrandmasters ($SE = 9.066, p < .001$) and candidates and nongrandmasters ($SE = 26.970, p < .001$) did. The next smallest difference between candidates and noncandidate grandmasters occurred in the 100–149 games category, and all pairs of groups differed significantly here. The overall ANOVA was highly significant $[F(2,324) = 89.094, p < .001]$. Candidates and noncandidate grandmasters differed significantly ($SE = 22.686, p < .01$), as did noncandidate grandmasters and nongrandmasters ($SE = 7.743, p < .001$) and candidates and nongrandmasters ($SE = 23.032, p < .001$).

An ANOVA for the three groups using ratings in the last four games categories (from 700–749 to 850–899 games) showed that these group differences then really were constant and the ratings were at asymptote. The

ANOVA showed a highly significant effect of group $[F(2,324) = 279.904, p < .001]$ and, with Greenhouse–Geisser correction, showed no significant effect of point $[F(1.556,504.227) = 0.079]$ and no significant interaction $[F(3.113,504.227) = 0.603]$.

Therefore, the interaction may be characterized as follows: Differences between candidates and nongrandmasters and between noncandidate grandmasters and nongrandmasters are present at all levels, but those between candidates and noncandidate grandmasters appear only by the 100–149 games category. Differences progressively widen, eventually staying at asymptote and constant from the 700–749 games category.

The averaged curves in Figure 4 are excellent fits to logarithmic and power functions. The best fit for all players was a logarithmic function (adjusted $R^2 = .991$), but the fit was only marginally better than power (adjusted $R^2 = .99$). For candidates, the best fit was logarithmic (adjusted $R^2 = .984$) but only marginally better than power (adjusted $R^2 = .983$). That for noncandidate grandmasters was best fitted by a logarithmic model (adjusted $R^2 = .994$) (for power, adjusted $R^2 = .993$). For nongrandmasters, the best fit was logarithmic (adjusted $R^2 = .973$) but only marginally better than for power (adjusted $R^2 = .972$).

Discussion

Overall expertise development patterns and individual differences for number of games are similar to those over time. Eventual top players are identifiable early and consistently perform much better than others, despite equivalent amounts of actual practice.

STUDY 5

Extreme Amounts of Actual Practice

Study 4 suggested that players approach rating asymptotes after about 750 games. However, only ratings up to the 850–899 games category were examined. Perhaps players would eventually improve further after playing many more games. Study 5 aimed to gauge the effects of much greater amounts of actual practice, more than 2,350

games, which represents probably more than 8,000 h of play. Do players continually improve with extreme amounts of actual practice or stay at a performance ceiling reached much earlier? If amount of actual practice is the only key factor, performance should keep rising. If, as proposed by Galton (1869/1979), natural talent places an immutable limit on performance, players should stay at an early performance ceiling.

Method

Participants. The participants were the 5 players who played the most games from July 1985 to October 2007. All entered the list before July 1985, but their game counts before then and precisely when they reached 850 FIDE-rated games are unknown. All played games in every year from 1986 on, and all played at least 89 games in 2007. All were male, and all were grandmasters. Their median age by October 2007 was 45 years (see Table 4).

Procedure. FIDE gave no rating for Player C for January 2002 or for Player E for January 1984, and these values were estimated by taking the mean in adjacent cells for each player. The final rating in each game category was determined for each player as in Study 4. Because these very active players often played more than 50 games in a rating period, about 19% of the values were missing. Missing values were estimated by taking the mean of ratings in adjacent categories. Players A and B had no value in the 0–49 games category, and their ratings in this category were not estimated. Group means were calculated from the first category in which all players had a rating (50–99 games), up until the final category in which all players had a value (2,350–2,399 games).

Results

Table 4 presents summary data on overall performance. Figures 5 and 6 present individual data in all game categories, and Figure 7 presents mean ratings from the 50–99 games category until the 2,350–2,399 category. Figures 5 and 6 suggest that the players had reached an approximate rating ceiling, with ratings fluctuating only within a 100-point range. Certainly, ratings did not improve with further practice for most of the span. Figure 7 suggests no upward trend from the peak value at the 1,150–1,199 games category. Indeed, an ANOVA on ratings in all categories from 850–899 games to 2,350–2,399 games with Greenhouse–

Table 4
Data of the Five Most Active Players From July 1985 to October 2007

Player	Total Games From July 1985	Age on Oct. 1, 2007	Age First on List	Peak Rating	Peak Rank
A	2,607	61	23	2540	44
B	2,570	45	19	2585	75
C	2,452	60	22	2635	18
D	2,362	42	18	2510	177
E	2,355	42	14	2675	11

Note—All players entered the list before July 1985.

Geisser correction showed no significant effect of games category [$F(3.683, 14.731) = 1.777, p < .2$].

A least squares analysis on the five individual curves (using all their data from 850–899 games onward) and on the mean value for all 5 players in each category (from 850–899 games to 2,350–2,399 games) compared power, linear, logarithmic, growth, and exponential models. The averaged curve best fit a linear model (adjusted $R^2 = .38$). Adjusted R^2 for the logarithmic model was .293, and for power, .294. Individual curves fit no consistent pattern. Only Player E showed a best fit to a logarithmic curve, and it was a poor one (adjusted $R^2 = .12$). The linear model was the best fit for Player A (adjusted $R^2 = .688$) and Player B (adjusted $R^2 = .092$) and equal best fit for Player C (adjusted $R^2 = .006$ for linear, exponential, and growth). For Player D, growth and exponential curves fit equally well (adjusted $R^2 = .059$).

Discussion

Even extreme amounts of additional actual practice did not push the players beyond an approximate asymptote reached early on. Despite extensive further actual practice, they improved no more. The most plausible account is that they had reached their natural talent limits, which were below those of candidates. However, one might argue that they were content with their peak performance levels and

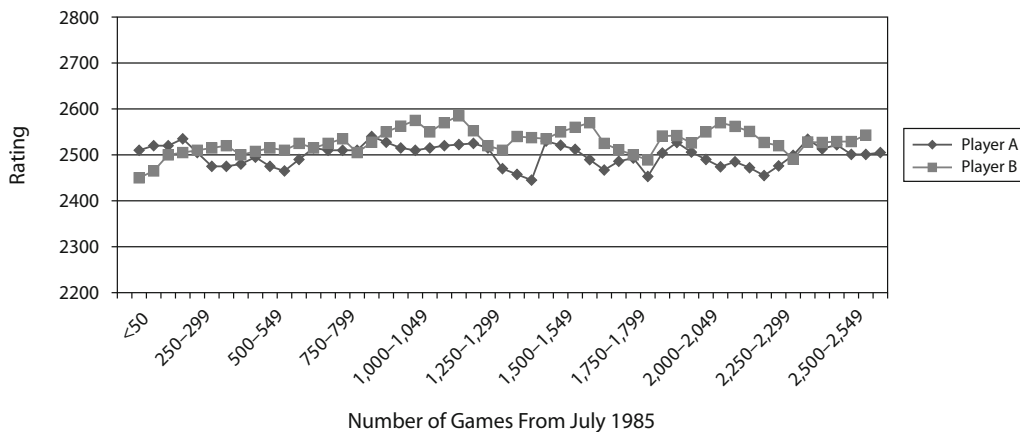


Figure 5. Ratings over cumulative number of games for 2 of the 5 players who played the most games from July 1985 through October 2007.

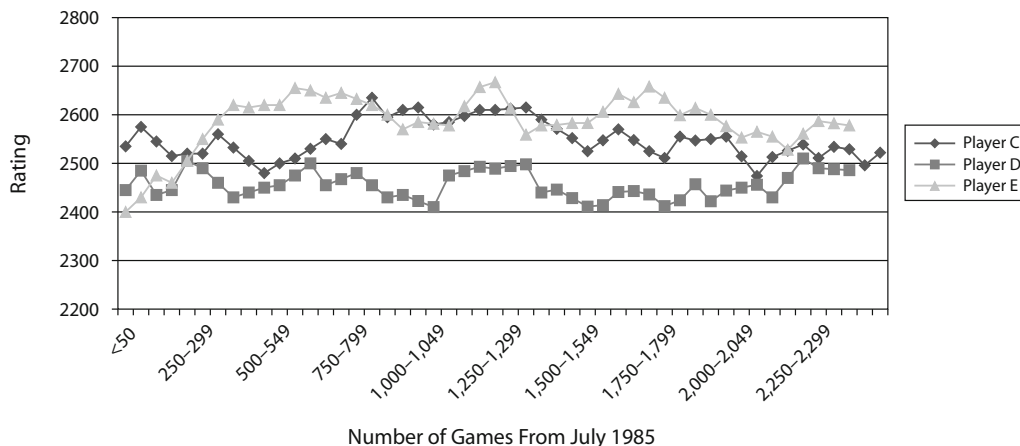


Figure 6. Ratings over cumulative number of games for 3 of the 5 players who played the most games from July 1985 through October 2007.

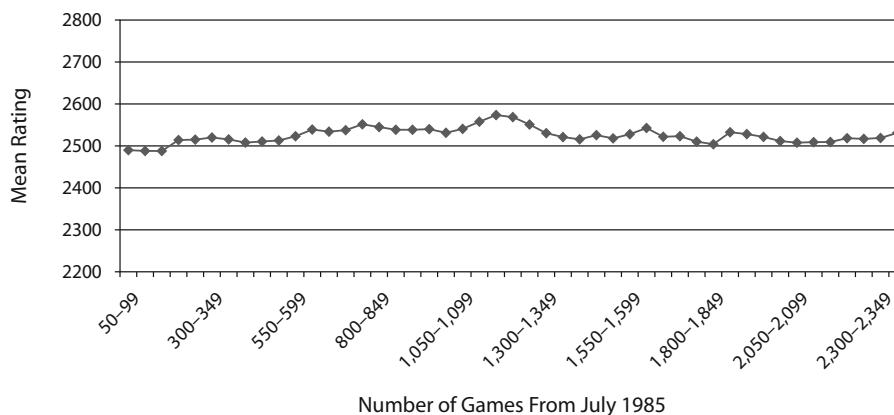


Figure 7. Mean ratings from 50–99 games to 2,350–2,399 games for the 5 players who played the most games from July 1985 through October 2007. All entered the list before July 1985.

did not undertake the difficult deliberate practice needed to rise further. There is no evidence for such complacency. Furthermore, such active players would be passionate about chess. They averaged more than three times as many games as did world first-ranked Kasparov in this period. If they believed that ascending further depended only on more deliberate practice, surely they would have undertaken it.

Another account is that cognitive aging might have masked possible continual improvement due to practice. However, this account is unlikely. The age of peak performance in chess has long been believed to be around 35 years old, but using the present FIDE data set, Roring and Charness (2007) found that the age of peak performance of international players is around 43.8 years old. Table 4 shows that 3 of the 5 participants were around this age by 2007.

GENERAL DISCUSSION

The major aim of the present study was to see what patterns of individual differences in expertise development

occur over decades and over extensive actual practice in a complex domain. The predominant expertise development pattern in grouped data was a power or logarithmic function, reaching asymptote after extensive actual practice. Fits to power functions from list entry were excellent, as the power law of practice would predict. However, in most cases, the fit to a logarithmic function was marginally better. Power and logarithmic functions are closely related, but perhaps the power law of practice might not be as ubiquitous as thought. Indeed, Heathcote et al. (2000) argued that an exponential function is often a better fit to practice and performance data than is a power function, though this was not found in the present study. Research might investigate this issue further.

The typical pattern of individual differences was an initial performance difference at domain entry between eventual top players and others; this difference widened with actual practice and time, reached asymptote after large amounts of practice, and then stayed relatively constant. This pattern contrasts with a common finding in the vocational

domain of initial differences, which tend to converge with practice (Ackerman, 2007). However, this pattern has been found elsewhere with very large amounts of practice—in mathematics education, for instance (Ackerman, 2007). Baltes and Kliegl (1992) compared word recall of younger and older participants who had extensive practice in use of the method of loci, a mnemonic device. Differences in number of words recalled between younger and older groups were small initially, but then widened and stayed large and constant. In science, there is anecdotal evidence that productivity differences between top performers and others occurring at career start may progressively widen.

Whereas the present pattern found in chess data might not be so unusual when large amounts of practice are involved, its causes may differ across domains. Poor understanding of chess material is an unlikely account of differences between candidates and noncandidate grandmasters. The latter would easily understand chess material. In science, a widening of an initial productivity difference between top and lesser producers may be largely magnified by gatekeeper influences and by occurrence of the “Matthew effect” (“To him that hath is given”) (Di Prete & Eirich, 2006; Ferrales & Fine, 2005; Simonton, 2008). Initially successful scientists may gravitate to universities with excellent research facilities and light teaching loads, find it much easier to get grants and to gather large research teams, and thus may become progressively more productive. Gatekeeper influences are small in chess, but cumulative advantage may partly explain the present pattern. For instance, initially successful players may get much more time to study chess (e.g., if they become professionals) and better coaching. However, research might investigate further why this pattern occurs in chess.

The present study suggests several phenomena of interest to researchers who study expertise. First, on average, players in Study 4 approached peak performance levels after around 750 rated games, regardless of whatever deliberate practice they may have had before or after list entry, how many nonrated games they may have played, and, evidently, what natural talent they had. Study 5 revealed no further improvement after extensive actual practice. However, performance levels at approximate asymptote differed greatly. One interpretation of these findings is that players reach the limits of their natural talent, as is consistent with anecdotal evidence in chess and with Galton’s (1869/1979) notion of differing immutable performance limits. Another interpretation of these findings might be that players have motivational differences—that is, they remain content with that asymptotic level of performance. However, the latter seems implausible with regard to top international players. It seems unlikely that they would all reach this decision at the same time, unless perhaps they all then realized their natural talent limits. Also, the asymptotic level correlates with other expected signs of natural talent, such as precocity and speed of expertise acquisition, again as predicted by the natural talent account.

Second, attrition rates over number of games are high, even with expert players rated over 2200. Most players do

not get enough actual practice to have a realistic chance of becoming a grandmaster. Expertise researchers who undertake retrospective studies should note possible high attrition rates. Comparisons between practiced experts and less practiced nonexperts may suggest multiple interpretations. Ideally, for researchers to study individual differences fully, all participants should eventually be very well practiced and near asymptote. Practice should proceed to performance limits, if possible (Baltes, 1998).

Another aim of the present study was to try to cast light on the issue of natural talent. Although this study is correlational, it allows, for the first time, a view of events within an entire population in an important domain, with objective measures of actual practice, performance, and dropout. Several specific predictions of the natural talent view were tested, and these predictions were upheld. All indicators suggest that, as expected, natural talent is implicated early and consistently and limits ultimate performance level. The patterns found in Studies 3 and 4 support Galton’s (1869/1979) argument about natural talent limits, as does the apparent inability of players in Study 5 to reach the top 10. The present results for chess partly parallel those found in various other intellectual areas. In science and the arts, for instance, most persons who contribute anything contribute at best perhaps one work a year. However, a few “superproducers” make many more contributions and show signs of precocity. The performance contrast between eventual superproducers and less productive performers appears early, at the beginning of one’s career, and it continues on to the career’s termination (Simonton, 1997, 2007). Simonton (2008) has noted that a good predictor of eventual success in science is the number of publications a young scientist makes before earning a doctorate. Such precocity predicts eventual high performance in science, just as in chess.

The present study could be criticized on several grounds, however. First, one might target the FIDE data and the study’s correlational nature. The FIDE data are not perfect but have huge advantages: They are objective, include everyone in the domain, and allow a look at expertise development over very long periods. Are the FIDE game counts a good measure of actual practice? They are an excellent measure of actual practice at playing FIDE-rated games. The FIDE game counts do not include offhand and nonrated games, but it is unclear how much these might contribute to FIDE ratings. The approach to asymptotes at around the same time suggests that differences in other forms of practice are relatively unimportant. Regardless, nonrated games could not be measured with this data set, and FIDE ratings measure skill at playing FIDE-rated games, not at playing offhand or speeded games.

Second, one might argue that the results are due solely to differences in amount of and quality of deliberate practice. The problems of defining and measuring deliberate practice in retrospective, correlational studies were noted in the Evidence on the Natural Talent Issue section. A deliberate practice account could explain any pattern of results with additional assumptions about amounts and quality of practice or about the need to start very young.

The natural talent view could readily explain only one pattern: the exact pattern that occurred. Deliberate practice could not be measured with the present data set, and any claims that the results are due to deliberate practice are essentially untestable.

Third, age was not controlled for. Perhaps cognitive aging might have masked continual improvement due to practice. However, this account is unlikely. Different fields have different ages of peak performance and different career landmarks (Simonton, 1997, 2007), but the peak age in chess is quite late. As mentioned, this account is an unlikely explanation for the Study 5 data. For the Study 3 and 4 data, this account is also an unlikely explanation, even if we consider peak age to be 35 years old. Table 1 shows that grandmasters and nongrandmasters who had played 740 games or more were first on the list on average before the age of 20; the candidates were about 6 years younger on average than the nongrandmasters. The asymptotic rating level occurred on average about 12 years after list entry, so, on average, no group had reached its peak age by then. In fact, the groups that got on at older median ages were closer to their peak ages and should have had an advantage over the candidates, all other factors being equal. But the results do not support such an advantage.

The present study provides evidence for the existence of natural talent for chess, but not about its nature. Probably many ability and personality characteristics are involved, as is apparently the case with talent for science and other fields (Simonton, 2008). As mentioned, research suggests that the top performers in a domain (superproducers) may differ in abilities and “have identifiably different character traits” than do run-of-the-mill experts (Simonton, 2000, p. 285). These traits include independence, aggressiveness, risk-taking, nonconformity, a broad range of interests, and greater versatility. Some research suggests that top athletes in physical sports also have certain psychological traits, such as competitiveness, mental toughness, and resilience (Gould, Dieffenbach, & Moffett, 2002).

As mentioned, some aspects of great natural talent for chess may include a minimum IQ of 120, good visuospatial ability, and creativity. Another might be an excellent memory, which is needed to acquire and continually update huge amounts of chess material. *Der Spiegel* (1987) found that Garry Kasparov had a reasonably high IQ score and an excellent memory. Other factors might include high motivation and extreme competitiveness. In a study of Dutch players, de Bruin, Rikers, and Schmidt (2007) suggested that high motivation to study chess material is important. Russian grandmaster Nikolai Krogius (1976) suggested the importance of ability to focus and maintain attention, to control emotions, to manage time well, and to analyze one's own strengths and weaknesses. However, further research is needed. One avenue would be to determine whether candidates and the very well-practiced nongrandmasters in the present study differ on such ability and personality factors.

Finally, whatever implications the present study has for the issue of natural talent, the data reveal several interesting phenomena of expertise development. And here, for the first time, there is a clear view of patterns of individual

differences in expertise development over very long intervals and amounts of actual practice in a domain with no sampling problems and little gatekeeper influence.

AUTHOR NOTE

Send correspondence to R. W. Howard, School of Education, University of New South Wales, Sydney 2052, Australia (e-mail: rwh@unsw.edu.au).

REFERENCES

- ACKERMAN, P. L. (2007). New developments in understanding skilled performance. *Current Directions in Psychological Science*, *16*, 235-239.
- ACKERMAN, P. L., & BEIER, M. E. (2006). Methods for studying the structure of expertise: Psychometric approaches. In K. A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 147-165). Cambridge: Cambridge University Press.
- ACKERMAN, P. L., & CIANCIOLO, A. T. (2000). Cognitive, perceptual-speed, and psychomotor determinants of individual differences during skill acquisition. *Journal of Experimental Psychology: Applied*, *6*, 259-290.
- BALTES, P. B. (1998). Testing the limits of the ontogenetic sources of talent and excellence. *Behavioral & Brain Sciences*, *21*, 407-408.
- BALTES, P. B., & KLIIEGL, R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology*, *28*, 121-125.
- BARRON, F., & HARRINGTON, D. M. (1981). Creativity, intelligence, and personality. *Annual Review of Psychology*, *32*, 439-476.
- BILALIC, M., MCLEOD, P., & GOBET, F. (2007). Does chess need intelligence?—A study with young chess players. *Intelligence*, *35*, 457-470.
- 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, sports, and games* (pp. 51-80). Mahwah, NJ: Erlbaum.
- 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.
- CHASE, W. G., & SIMON, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.
- DEAKIN, J. M., COTE, J., & HARVEY, A. S. (2006). Time budgets, diaries, and analyses of concurrent practice activities. In K. A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 303-318). Cambridge: Cambridge University Press.
- 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*, 561-583.
- DETERMAN, D. K., GABRIEL, L. T., & RUTHSATZ, J. M. (1998). Absurd environmentalism. *Behavioral & Brain Sciences*, *21*, 411-412.
- DETERMAN, D. K., & SPRY, K. M. (1988). Is it smart to play the horses? Comment on “A day at the races: A study of IQ, expertise and cognitive complexity” (Ceci & Liker, 1986). *Journal of Experimental Psychology: General*, *117*, 91-95.
- DI PRETE, T. A., & EIRICH, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology*, *32*, 271-297.
- DIJAKOV, I. N., PETROWSKI, N. W., & RUDIK, P. A. (1927). *Psychologie des Schachspiels*. Berlin: de Gruyter.
- DOLL, J., & MAYR, U. (1987). Intelligenz und Schachleistung—Eine Untersuchung an Schachexperten. *Psychologische Beiträge*, *29*, 270-289.
- DOWNY, J. E., & ANDERSON, J. E. (1917). Retention of skill after lapse of practice: Simultaneous reading and writing. *American Journal of Psychology*, *28*, 396-408.
- ELO, A. E. (1986). *The rating of chess players, past and present* (2nd ed.). New York: Arco.
- ERICSSON, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A.

- Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683-703). Cambridge: Cambridge University Press.
- ERICSSON, K. A., & CHARNES, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, **49**, 725-747.
- ERICSSON, K. A., NANDAGOPAL, K., & RORING, R. W. (2005). Giftedness viewed from the expert-performance perspective. *Journal for the Education of the Gifted*, **28**, 287-311.
- 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.
- ERICSSON, K. A., & WARD, P. (2007). Capturing the naturally occurring superior performance of experts in the laboratory: Toward a science of expert and exceptional performance. *Current Directions in Psychological Science*, **16**, 346-350.
- FERRALES, G., & FINE, G. A. (2005). Sociology as a vocation: Reputations and group cultures in graduate school. *American Sociologist*, **36**, 57-75.
- FRYDMAN, M., & LYNN, R. (1992). The general intelligence and spatial abilities of gifted young Belgian chess players. *British Journal of Psychology*, **83**, 233-235.
- GAGNÉ, F. (2007). Predictably, an unconvincing second attempt. *High Ability Studies*, **18**, 67-69.
- GALTON, F. (1979). *Hereditary genius: An inquiry into its laws and consequences*. London: Julian Freeman. (Original work published 1869)
- GARDNER, H. (1995). "Expert performance: Its structure and acquisition": Comment. *American Psychologist*, **50**, 802-803.
- GOULD, D., DIEFFENBACH, K., & MOFFETT, A. (2002). Psychological characteristics and their development in Olympic champions. *Journal of Applied Sport Psychology*, **14**, 172-204.
- GRABNER, R. H., NEUBAUER, A. C., & STERN, E. (2006). Superior performance and neural efficiency: The impact of intelligence and expertise. *Brain Research Bulletin*, **69**, 422-439.
- GRABNER, R. H., STERN, E., & NEUBAUER, A. C. (2007). Individual differences in chess expertise: A psychometric investigation. *Acta Psychologica*, **124**, 398-420.
- GREGORY, R. J. (2004). *Psychological testing: History, principles, and applications*. Needham Heights, MA: Allyn & Bacon.
- HEATHCOTE, A., BROWN, S., & MEWHORT, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, **7**, 185-207.
- HODGES, N. J., & STARKES, J. L. (1996). Wrestling with the nature of expertise: A sport specific test of Ericsson, Krampe and Tesch-Römer's (1993) theory of "deliberate practice." *International Journal of Sport Psychology*, **27**, 400-424.
- HORGAN, D. D., & MORGAN, D. (1990). Chess expertise in children. *Applied Cognitive Psychology*, **4**, 109-128.
- HOWARD, R. W. (1995). *Learning and memory: Major ideas, principles, issues and applications*. Westport, CT: Praeger.
- HOWARD, R. W. (2005). Objective evidence of rising population ability: A detailed examination of longitudinal chess data. *Personality & Individual Differences*, **38**, 347-363.
- HOWARD, R. W. (2006). A complete database of international chess players and performance ratings for varied longitudinal studies. *Behavior Research Methods*, **38**, 698-703.
- HOWE, M. J. A., DAVIDSON, J. W., & SLOBODA, J. A. (1998). Innate talents: Reality or myth. *Behavioral & Brain Sciences*, **21**, 399-442.
- HUNT, E. (1995). The role of intelligence in modern society. *American Scientist*, **83**, 356-368.
- HUNT, E. (2006). Expertise, talent and social encouragement. In K. A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 31-38). Cambridge: Cambridge University Press.
- KIM, K. H. (2005). Can only intelligent people be creative? A meta-analysis. *Journal of Secondary Gifted Education*, **16**, 57-66.
- KROGIUS, N. V. (1976). *Psychology in chess*. Albertson, NY: RHM Press.
- MCINTIRE, S. A., & MILLER, L. A. (2000). *Foundations of psychological testing: A practical approach*. New York: McGraw-Hill.
- NEWELL, A., & ROSENBLUM, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-56). Hillsdale, NJ: Erlbaum.
- REE, M. J., & EARLES, M. (1991). Predicting training success: Not much more than g. *Personnel Psychology*, **44**, 321-332.
- RORING, R. W., & CHARNES, N. (2007). A multilevel model analysis of expertise in chess across the life span. *Psychology & Aging*, **22**, 291-299.
- ROSS, P. E. (2006). The expert mind. *Scientific American*, **295**, 64-71.
- SCHMIDT, F. L., & HUNTER, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, **124**, 262-274.
- SHRAGER, J., HOGG, T., & HUBERMAN, B. A. (1988). A graph-dynamic model of the power law of practice and the problem-solving fan-effect. *Science*, **242**, 414-416.
- SIMONTON, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review*, **104**, 66-89.
- SIMONTON, D. K. (1999). Talent and its development: An emergent and epigenetic model. *Psychological Review*, **106**, 435-457.
- SIMONTON, D. K. (2000). Creative development as acquired expertise: Theoretical issues and an empirical test. *Developmental Review*, **20**, 283-318.
- SIMONTON, D. K. (2007). Cinema composers: Career trajectories for creative productivity in film music. *Psychology of Aesthetics, Creativity, & the Arts*, **1**, 160-169.
- SIMONTON, D. K. (2008). Scientific talent, training and performance: Intellect, personality, and genetic endowment. *Review of General Psychology*, **12**, 28-46.
- SMYTH, M. M. (1975). The role of mental practice in skill acquisition. *Journal of Motor Behavior*, **7**, 199-206.
- DER SPIEGEL (1987). Genieblitze und Blackouts: Der Spiegel testete Intelligenz, Gedächtnis, und Schachkunst Garri Kasparows. [Strokes of genius and blackouts: Der Spiegel tests the intelligence, memory, and chess ability of Garry Kasparov]. *Der Spiegel*, **41**, 126-140.
- STERNBERG, R. J. (1996). Costs of expertise. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 347-354). Mahwah, NJ: Erlbaum.
- TALWAR, G. P., & SRIVASTAVA, L. M. (2004). *Textbook of biochemistry and human biology* (3rd ed.). New Delhi: PHI Learning.
- WARGO, E. (2006, August). The myth of prodigy and why it matters. *APS Observer*, **19**. Retrieved November 18, 2008, from www.psychologicalscience.org/observer/getArticle.cfm?id=2026.
- WATERS, A. J., GOBET, F., & LEYDEN, G. (2002). Visuospatial abilities of chess players. *British Journal of Psychology*, **93**, 557-565.

(Manuscript received June 13, 2008;
revision accepted for publication October 10, 2008.)