Intelligence 45 (2014) 34–45

Contents lists available at ScienceDirect

Intelligence

Deliberate practice: Is that all it takes to become an expert?☆

David Z. Hambrick [a,*], Frederick L. Oswald [b], Erik M. Altmann [a], Elizabeth J. Meinz [c], Fernand Gobet [d], Guillermo Campitelli [e]

[a] *Department of Psychology, Michigan State University, United States*
[b] *Department of Psychology, Rice University, United States*
[c] *Department of Psychology, Southern Illinois University Edwardsville, United States*
[d] *Department of Psychology, Brunel University, United Kingdom*
[e] *School of Psychology and Social Science, Edith Cowan University, Australia*

ARTICLE INFO

Article history:
Received 7 February 2013
Received in revised form 15 March 2013
Accepted 5 April 2013
Available online 15 May 2013

Keywords:
Expert performance
Expertise
Deliberate practice
Talent

ABSTRACT

Twenty years ago, Ericsson, Krampe, and Tesch-Römer (1993) proposed that expert performance reflects a long period of *deliberate practice* rather than innate ability, or "talent". Ericsson et al. found that elite musicians had accumulated thousands of hours more deliberate practice than less accomplished musicians, and concluded that their theoretical framework could provide "a sufficient account of the major facts about the nature and scarcity of exceptional performance" (p. 392). The deliberate practice view has since gained popularity as a theoretical account of expert performance, but here we show that deliberate practice is not sufficient to explain individual differences in performance in the two most widely studied domains in expertise research—chess and music. For researchers interested in advancing the science of expert performance, the task now is to develop and rigorously test theories that take into account as many potentially relevant explanatory constructs as possible.

1. Introduction

Psychologists have been interested in how people become experts in music, the arts, science, games, sports, and professions for as long as psychology has been a field. Sir Francis Galton (1869) analyzed genealogical records of scientists, musicians, writers, poets, painters, athletes, and other men of "eminence" and found that they tended to be biologically related. He noted, for example, that there were over twenty eminent musicians in the Bach family. Acknowledging a role for "zeal" and "an adequate power of doing a great deal of very laborious work" (p. 37), Galton nonetheless concluded that "genius" arises from innate ability. John Watson (1930), the founder of behaviorism, famously captured the opposing view. Watson wrote:

> Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select—doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents… (p. 104).

Watson added that "practicing more intensively than others…is probably the most reasonable explanation we have today not only for success in any line, but even for genius" (p. 212). Thus the pendulum has swung between nature and nurture—the view that experts are "born" and the view that they are "made."

More recently, K. Anders Ericsson and his colleagues (Ericsson et al., 1993) sided with Watson when they proposed that expert performance–consistently performing at a superior

☆ This work is a collaboration between two research groups: the first through fourth authors, and the fifth and sixth authors. We are very grateful to David Lane, Lauren Harris, and Tom Carr for comments, and to Ramey Soukup for assistance in preparing the manuscript. We thank Rashika Vankataraman and staff for their care in copy editing and typesetting the article.

* Corresponding author at: Department of Psychology, Michigan State University, East Lansing, MI 48824, United States.
E-mail address: hambric3@msu.edu (D.Z. Hambrick).

0160-2896/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.intell.2013.04.001

level in a domain–reflects a long period of *deliberate practice* rather than innate ability, or "talent". Ericsson et al. defined deliberate practice as engagement in highly structured activities that are created specifically to improve performance in a domain through immediate feedback, that require a high level of concentration, and that are not inherently enjoyable. Ericsson et al. distinguished deliberate practice from two other forms of domain-specific experience–*work* and *play*–as follows:

> Work includes public performance, competitions, services rendered for pay, and other activities directly motivated by external rewards. Play includes activities that have no explicit goal and that are inherently enjoyable. Deliberate practice includes activities that have been specially designed to improve the current level of performance (p. 368).

To test their theory, Ericsson et al. (1993) asked violinists at a West Berlin music academy to rate various activities on relevance to improving violin performance, on effort, and on enjoyableness. They also asked the violinists to provide estimates of the time they devoted to the activities. Ericsson et al. found that the students whom the faculty had nominated as the "best" violinists had accumulated an average of over 10,000 h of deliberate practice by age 20, which was about 2500 h more than the average for the "good" violinists and about 5000 h more than the average for a "teacher" group (see Ericsson, 2006). In a second study, Ericsson et al. (1993) found that "expert" pianists had similarly accumulated an average of over 10,000 h of deliberate practice by age 20, compared to about 2000 h for "amateur" pianists.

Applying their framework to several domains, Ericsson et al. (1993) concluded that "high levels of deliberate practice are necessary to attain expert level performance" (p. 392)—and in the next sentence added the following:

> Our theoretical framework can *also provide a sufficient account* of the major facts about the nature and scarcity of exceptional performance. Our account does not depend on scarcity of innate ability (talent)…. We attribute the dramatic differences in performance between experts and amateurs–novices to similarly large differences in the recorded amounts of deliberate practice (p. 392, emphasis added).

Ericsson et al. (1993) similarly explained that "individual differences in ultimate performance can largely be accounted for by differential amounts of past and current levels of practice" (p. 392) and that "the differences between expert performers and normal adults reflect a life-long period of deliberate effort to improve performance in a specific domain" (p. 400).

Ericsson et al. (1993) allowed that genes may contribute to individual differences in people's willingness to engage in deliberate practice over a long period of time, and thus may indirectly contribute to individual differences in performance, but as the preceding quotations make clear, they explicitly rejected the view that innate *ability* can account for why some people become experts and others fail to do so. Ericsson, Nandagopal, and Roring (2005) recently reiterated this perspective when they wrote that

> individual differences in genetically determined capacities and fixed structures required for the development of elite performance appear to be quite limited, perhaps even restricted, to a small number of physical characteristics, such as height and body size. The expert-performance framework attempts to explain the large individual differences in performance in terms of individual differences in sustained deliberate practice (p. 305).

Similarly, Ericsson (2007) argued that "it is possible to account for the development of elite performance among healthy children without recourse to unique talent (genetic endowment)—excepting the innate determinants of body size" (p. 4) and that "distinctive characteristics of elite performers are adaptations to extended and intense practice activities that selectively activate dormant genes that all healthy children's DNA contain" (p. 4). Ericsson, Prietula, and Cokely (2007) wrote more simply that "The *only* innate differences that turn out to be significant–and they matter primarily in sports–are height and body size" (p. 116, emphasis added).

1.1. Impact and criticisms of the deliberate practice view

As two of us noted in a recent *New York Times* op-ed (Hambrick & Meinz, 2011a), Ericsson and colleagues' research has captured the popular imagination, perhaps because of its meritocratic appeal—the implication that nearly anyone can become an expert with enough hard work. In his bestselling book *Outliers*, the writer Malcolm Gladwell (2008) summarized Ericsson et al.'s (1993) findings and then described the opportunity to practice as the major theme of the biographies of Bill Gates and The Beatles. Ericsson and colleagues' research is discussed in a number of other popular books, including Daniel Levitin's (2006) *This is Your Brain on Music*, Geoff Colvin's (2010) *Talent is Overrated*, Daniel Pink's (2009) Drive, Daniel Coyle's (2009) *The Talent Code*, David Shenk's (2010) *The Genius in All of Us*, Matthew Syed's (2010) *Bounce*, and David Brooks' (2011) *The Social Animal*.

The Ericsson et al. (1993) article has been cited in the scientific literature over a thousand times (source: Web of Science), making it a "citation classic" many times over, and Ericsson and colleagues have been praised for advancing scientific understanding of expert performance. Freeman (2007) observed that "The field of gifted and talented research is in serious need of scientific work of this calibre, as distinct from theories, models and anecdotes" (p. 65), and Kaufman (2007) commented that "The expert performance approach championed by Ericsson et al. provides a scientific way forward for research on giftedness, and offers exciting new ways to further our understanding of the determinants of high ability within a particular domain of expertise" (p. 71).

At the same time, Ericsson and colleagues' view has been roundly criticized on conceptual and methodological grounds. Gardner (1995) commented that the deliberate practice view "requires a blindness to ordinary experience" (p. 802), and Sternberg (1996) observed that "Most people who want to become experts–whether as violinists, skiers, physicists, or whatever–do not make it. They drop out along the way" (p. 350). Schneider (1998) questioned "the basic assumption that progress in a given domain is solely a function of deliberate practice" (p. 424), and Detterman, Gabriel, and Ruthsatz (1998) predicted that deliberate practice "will not equalize outcome despite the best of intentions" (p. 412). Anderson

(2000) concluded that "Ericsson and Krampe's research does not really establish the case that a great deal of practice is sufficient for great talent" (p. 324), and Winner (2000) observed that "Ericsson's research demonstrated the importance of hard work but did not rule out the role of innate ability" (p. 160). Marcus (2012) noted that deliberate practice "does indeed matter–a lot–and in surprising ways" but that "it would be a logical error to infer from the importance of practice that talent is somehow irrelevant, as if the two were in mutual opposition" (p. 94), and Ackerman (2014) added that factors other than deliberate practice "must clearly play a role in the demonstration of expert performance" (p. 11). Finally, both Hambrick and Meinz (2011b) and Campitelli and Gobet (2011) concluded that deliberate practice is necessary but not sufficient to account for individual differences in performance. There is widespread skepticism, then, over Ericsson and colleagues' strong claims regarding the importance of deliberate practice for acquiring expert performance.

2. Present research

Here, we evaluated the deliberate practice view on empirical grounds. The fundamental question that we set out to address is whether deliberate practice is as important for explaining individual differences in performance as Ericsson and colleagues have argued it is. That is, can individual differences in performance largely be accounted for by individual differences in deliberate practice? Is deliberate practice essentially all it takes to become an expert?

To answer this question, we reanalyzed findings from research on the two most widely studied domains in expertise research: chess and music. There were two criteria for including a study in the reanalysis: (a) continuous measures of performance and of cumulative amount of time engaged in activity interpretable as deliberate practice were collected, and (b) a correlation between these measures was reported. For a given study, our question was how much of the variance in performance deliberate practice explained.

To foreshadow, we found that deliberate practice does not account for all, nearly all, or even most of the variance in performance in these domains. We conclude that deliberate practice is not nearly as important for explaining individual differences in performance as Ericsson and colleagues have argued it is, and review evidence for other factors that may also directly contribute.

2.1. General method

As Spearman (1904) first observed, it is not possible to accurately test the magnitude of a correlation without controlling for the distorting effect of measurement error variance—that is, the unreliability of the measures (see Schmidt & Hunter, 1999, for an excellent review). This critical point is reflected in the formula in classical measurement theory for a correlation between two measures, x and y: $r_{xy} = r_{x_t y_t}\,(r_{xx} r_{yy})^{1/2}$, where r_{xy} is the observed correlation between the measures, $r_{x_t y_t}$ is the true-score correlation reflecting the constructs underlying the measures, and r_{xx} and r_{yy} are the reliabilities of the measures (Hunter & Schmidt, 1990). As Schmidt and Hunter (1999) explain, this formula is called the *attenuation formula* because it demonstrates that measurement error variance (i.e., unreliability) in one or both measures reduces the observed correlation relative to the true-score correlation. Solving algebraically for $r_{x_t y_t}$ yields the standard *disattenuation formula* for correcting correlations for measurement error variance: $r_{x_t y_t} = r_{xy} \,/\, (r_{xx} r_{yy})^{1/2}$. We used the disattenuation formula to correct correlations for measurement error variance in both deliberate practice and performance. For a given study, we report the observed correlation and corrected correlation ($\hat{r}$), with 95% confidence intervals.

The reliability of retrospective estimates of deliberate practice is of particular concern, because people obviously do not have perfect memory for the past. However, Tuffiash, Roring, and Ericsson (2007) stated that "self-report practice estimates repeatedly from experts in sports and music have reported test–retest reliabilities at or above .80" (p. 129), which is "good" or better reliability by psychometric standards (Cronbach, 1990). Similarly, Ericsson (2012a) claimed that "The collected reliability of cumulated life-time practice at different test occasions in large samples has typically been found to range between 0.7 and 0.8" (p. 534). Indeed, several studies have demonstrated that estimates of cumulative deliberate practice have reasonably high reliability. Ericsson et al. (1993) found a correlation of .74 between retrospective and diary-based estimates of deliberate practice for a current typical week and concluded that "subjects should be able to accurately report not just their current level of practice, but past levels of practice as well" (p. 379), and in a study of musicians, Tuffiash (2002) found a test–retest correlation of .89 between estimates of cumulative deliberate practice. de Bruin, Smits, Rikers, and Schmidt (2008) found a correlation of .60 between diary-based and retrospective estimates of practice (see also de Bruin, Rikers, & Schmidt's, 2007, report of this study), and Bilalić, McLeod, and Gobet (2007) found correlations of .99 and .98 between diary-based and retrospective estimates of chess practice over two school terms.

Based on this evidence, we assumed reliability of .80 for cumulative deliberate practice, in line with Ericsson and colleagues' estimates (Ericsson, 2012a; Tuffiash et al., 2007), but we also tested the sensitivity of our analyses to different reliability assumptions. We discuss the reliability of measures of chess performance and music performance separately by domain below.

3. Chess studies

Chess has been dubbed the *Drosophila*–the model "organism"–for research on expertise because it offers an objective measure of performance in the Elo rating, and because it is possible to develop representative tasks for laboratory research (e.g., choose-a-move tasks). We identified six studies, listed in Table 1, that met the criteria for our reanalysis.[1] Except where noted, participants in these studies completed a survey to assess chess-related experience in which they provided information such as the age they started playing chess seriously

[1] One study we did not include in our reanalysis was by Grabner, Stern, and Neubauer (2007). They collected measures of deliberate practice and chess rating in 90 chess players, which correlated .08 (*ns*). We omit this study, because the measure of deliberate practice only reflected the current year.

Table 1
Correlations between deliberate practice and chess performance.

Study	Sample	Chess rating M (SD)	N	r (95% CI)	$\hat{r}$ (95% CI)	% Var.
Charness et al. (2005) — Study 1	International; intermediate to grandmaster (adult)	2032 (278)	200	.54 (.43, .63)	.63 (.50, .74)	39.7 (60.3)
Charness et al. (2005) — Study 2	International; intermediate to grandmaster (adult)	2008 (253)	164	.48 (.35, .59)	.56 (.41, .69)	31.4 (68.6)
Gobet and Campitelli (2007)	Argentine; intermediate to grandmaster (adult)	1988 (209)	90	.42 (.23, .58)	.49 (.27, .68)	24.0 (76.0)
Bilalić et al. (2007)	British; intermediate (youth)	1603 (109)	23	.69 (.39, .86)	.81 (.46, 1.0)	65.6 (34.4)
de Bruin et al. (2008)	Dutch; intermediate to grandmaster (adolescent)	1944 (259)	73	.45 (.25, .62)	.53 (.29, .73)	28.1 (71.9)
Howard (2012)	International; intermediate to grandmaster (adult)	2122 (168)	533	.33 (.25, .40)	.39 (.29, .47)	15.2 (84.8)

Note. Studies are in chronological order. r, correlation between cumulative deliberate practice and chess performance. $\hat{r}$, corrected correlation. % var., percentage of variance in chess performance explained by deliberate practice (vs. unexplained) after correction for measurement error variance ($\hat{r}^2$). Reliability for deliberate practice (r_{xx}), .80. Reliability for chess rating (r_{yy}), .91. For Bilalić et al., the upper-bound of the 95% CI exceeds 1.0 and is truncated. Cumulative deliberate practice was log-transformed prior to analyses in all studies, except de Bruin et al. (2008).

and the number of years of instruction, and for each age since taking up the game estimated the number of hours per week they had spent *studying chess alone*—the activity that expertise researchers have identified as meeting the theoretical description of deliberate practice for chess (e.g., Ericsson, 2007; Ericsson & Charness, 1994). In each study, the measure that can be interpreted as reflecting cumulative deliberate practice was obtained by summing across annual estimates (i.e., hours per week × 52); except in de Bruin et al. (2008), this measure was log-transformed to normalize the data and allow for a linear test of the relationship between deliberate practice and performance. The measure of performance was chess rating, which the authors of the studies verified against published rating lists whenever possible.

Charness, Tuffiash, Krampe, Reingold, and Vasyukova's (2005) participants were chess players recruited from Canada, Germany, Russia, and the United States; in Study 1 there was no restriction on participation in terms of chess rating, whereas in Study 2 the minimum chess rating was 1600. Bilalić et al.'s (2007) participants included youth chess players. (The measure of practice in this study was the amount of time spent playing chess.) Gobet and Campitelli's (2007) participants were Argentine chess players recruited from a prestigious chess club in Buenos Aires. de Bruin et al.'s (2008) participants were youth who had been selected for a national chess training program run by the Dutch Chess Federation. Howard's (2012) participants were chess players who had completed an online survey with two questions about chess practice, which was advertised on chess websites and was available in several languages. (The measure of practice in this study was simply studying chess, instead of studying chess *alone*.) The standard deviation (SD) for chess rating in these studies was in most cases close to the theoretical SD of 200 for chess rating (Elo, 1986) and the actual SD (of 239) for the January, 2012, World Chess Federation rating list.[2,3] The average SD across studies was 212.7, and thus there was no overall indication of range restriction in chess skill.

To obtain an estimate of the reliability of chess rating, we computed correlations between expected outcome in chess tournaments and actual outcome. We used the past five Chess Olympiads and the past five European Individual Championships. For the Chess Olympiads (Ns = 128 to 156), we correlated team rating (i.e., the average rating of each teams' five members) with tournament outcome. Similarly, for the European Individual Championship (Ns = 303 to 402), we correlated player rating with tournament outcome. Consistent with Charness et al.'s (2005) comment that chess rating is highly reliable, and with reports of strong correlations between chess rating and other measures of chess performance (e.g., Bilalić et al., 2007; Burns, 2004; van der Maas & Wagenmakers, 2005), the correlations were strongly positive, ranging from .93 to .96 for the Chess Olympiads (avg. r = .95) and from .83 to .93 for the European Individual Championships (avg. r = .88). We used the overall average of .91 as the reliability coefficient for chess rating in our reanalysis.[4]

3.1. Results of chess reanalyses

Correlations between deliberate practice and chess performance, with 95% confidence intervals (CIs), are displayed in Table 1. On average, deliberate practice explained 34% of the variance in performance after correcting for measurement error variance (avg. $\hat{r}$ = .57; sample size-weighted avg. $\hat{r}$ = .49), leaving 66% of the variance unexplained and potentially explainable by other factors (see Fig. 1). The 95% CI included 1.0 in Bilalić et al. (2007), but the confidence interval was very wide ($\hat{r}$ = .81, 95% CI, .46, 1.0) due to a small sample size (N = 23).[5] (Note also that the correlation between practice and chess performance dropped from .69 to .60 after Bilalić et al. statistically controlled for IQ, which yields $\hat{r}$ = .70, 95% CI, .29, .95.)

There was a wide range of ages in some of the studies (e.g., Charness et al., 2005), which could represent a confound. To control for age differences, we computed partial correlations

[2] The sample in Charness et al.'s (2005) Study 1 included a smaller sample described in a chapter by Charness, Krampe, and Mayr (1996); we did not include this smaller sample in our reanalysis given that it was a subset of the later sample. Gobet and Campitelli (2007) distinguished between and measured two types of deliberate practice, individual practice and group practice; we include only the results for individual practice because of the focus on this activity in the expertise literature as a form of deliberate practice.

[3] We thank Robert W. Howard for sending us this information.

[4] We deleted players that abandoned the tournament three rounds before the end or earlier (at most 3% of players, and usually less than 1%). We chose the Chess Olympiad and European Individual Championship to compute the reliability estimate, because in other Open tournaments the rate of forfeit is typically much higher.

[5] Bilalić et al.'s (2007) total sample included both the "elite" subsample (n = 23) listed in Table 1 and a subsample of unrated players (n = 34). In the total sample (N = 57), Bilalić et al. found a very high correlation (r = .90) between practice and scores on a "chess test" that included questions about the rules of chess and chess problems. However, this correlation was inflated, because practice accounted for only 29% of the variance in scores on the test after Bilalić et al. statistically controlling for confounding factors (age, gender, IQ).

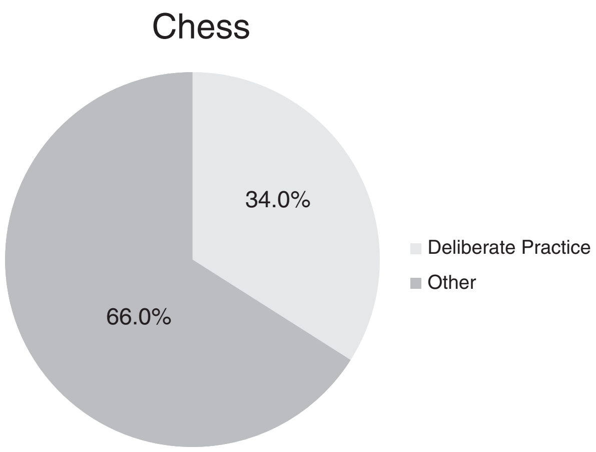

Fig. 1. Average percentage of variance in chess performance explained by deliberate practice, correcting for measurement error variance. The light gray region represents reliable variance explained by deliberate practice; the dark gray region represents reliable variance not explained by deliberate practice.

(*prs*) reflecting the relationship between deliberate practice and chess performance controlling for age. We then corrected the partial correlations for the unreliability of deliberate practice and chess rating (assuming reliability of .80 for deliberate practice and .91 for chess rating). The average corrected partial correlation was nearly the same as the average corrected correlation (avg. $p\hat{r} = .60$ vs. avg. $\hat{r} = .57$).

Our analyses would have underestimated the contribution of deliberate practice to performance if reliability of this measure was lower than .80, as it may have been in (for example) Howard's (2012) study given the use of a brief internet survey (see Ericsson & Moxley's, 2012, critique), or overestimated this contribution if it was higher than .80, as it may have been in Charness et al.'s (2005) exemplary studies. Therefore, we performed a sensitivity analysis in which we assumed different levels of reliability for deliberate practice. Results are shown in Table 2. Even assuming a low level of reliability ($r_{xx} = .60$), deliberate practice left a very large proportion of the variance in chess performance unexplained (54.7%).

3.2. Discussion of chess reanalyses

On average, deliberate practice explained 34% of the reliable variance in chess performance, leaving 66% unexplained and potentially explainable by other factors. We conclude that deliberate practice is not sufficient to account for individual differences in chess performance.

The implication of this conclusion is that some people require much less deliberate practice than other people to reach an elite level of performance in chess. We illustrate this point in Fig. 2 using Gobet and Campitelli's (2007) chess sample, with the 90 players classified based on their chess ratings as "master" ($\geq$2200, $n = 16$), "expert" ($\geq$2000, $n = 31$), or "intermediate" (<2000, $n = 43$). There were large differences in mean amount of deliberate practice across the skill groups: master $M =$ 10,530 h ($SD = 7414$), expert $M = 5673$ h ($SD = 4654$), and intermediate $M = 3179$ h ($SD = 4615$). However, as the *SDs* suggest, there were very large ranges of deliberate practice within skill groups. For example, the range for the masters was 832 to 24,284 h—*a difference of nearly three orders of magnitude*. Furthermore, there was overlap in distributions *between* skill groups. For example, of the 16 masters, 31.3% ($n = 5$) had less deliberate practice than the mean of the expert group, one skill level down, and 12.5% ($n = 2$) had less deliberate practice than the mean of the intermediate group, *two* skill levels down. In the other direction, of the 31 intermediates, 25.8% ($n = 8$) had more deliberate practice than the mean of the expert group, one skill level up, and 12.9% ($n = 4$) had more deliberate practice than the mean of the master group, *two* skill levels up.

Howard's (2011) case study of the three Polgár sisters provides further support for our conclusion. Beginning at a young age, the sisters received several hours of chess instruction every day from chess grandmasters and their father, a chess teacher and author of several chess books. Using practice estimates obtained from biographical and autobiographical accounts, Howard found that the sisters differed both in the highest rating they achieved and in the amount of practice they accumulated to reach that rating. For example, one sister's peak rating was 2735 in an estimated 59,904 h of practice, whereas another sister's was 2577–more than a standard deviation lower–in an estimated 79,248 h of practice. Howard also found that the two sisters who became grandmasters had accumulated a great deal more practice by the time they reached their peak rating than had the eight grandmasters in his sample who reached top-ten in the world ($M = 14{,}020.5$ h, $SD =$ 7373.96 h). Deliberate practice is clearly not sufficient to account for individual differences in chess performance.

4. Music

Music is another popular domain for psychological research on expertise. In fact, at least as many articles have been published on music expertise as on chess expertise (see Hambrick & Meinz, 2012, for a review). We identified eight studies, listed in Table 3, that met the criteria for our reanalysis.[6] Participants in these studies completed a survey to assess music-related experience in which they provided information such as the age they started playing music seriously and number of years of lessons, and for each age since taking up their instrument estimated the number of hours per week they had spent practicing (Ericsson et al.'s, 1993, interview procedure, or a very similar one, was used in each study). In each study, the measure that can be interpreted as reflecting cumulative deliberate practice was obtained by summing across annual

Table 2
Results of reliability analysis for chess performance.

	Reliability of deliberate practice (r_{xx})			
	$r_{xx} = .60$	$r_{xx} = .70$	$r_{xx} = .80$	$r_{xx} = .90$
Avg. % variance	45.3 (54.7)	38.7 (61.3)	34.0 (66.0)	30.2 (69.8)

Note. Avg. % variance, average percentage of variance in chess performance explained (vs. unexplained) by deliberate practice.

[6] One study we did not include in our reanalysis was by Bonneville-Roussy et al. (2011). They collected measures of deliberate practice and music skill in 187 classical musicians, which correlated .23 ($p < .001$). We omit this study because the measure of deliberate practice was not a direct time estimate but instead was based on Likert ratings of frequency of engagement in practice activities (e.g., "When I do my daily practice…I slowly repeat difficult excerpts").

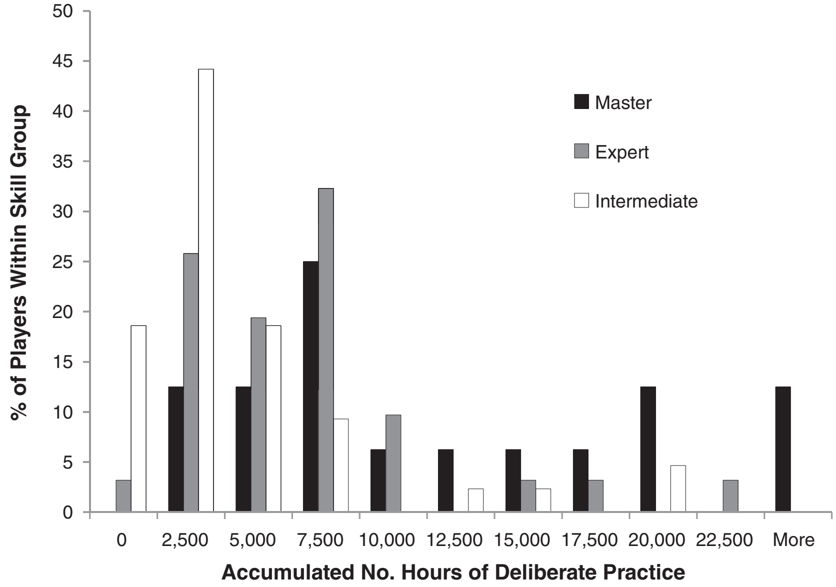

Fig. 2. Histograms showing accumulated hours of deliberate practice for "master" ($n = 16$), "expert" ($n = 31$), and "intermediate" ($n = 43$) chess players (Gobet & Campitelli, 2007). Deliberate practice refers to *study alone.*

estimates (i.e., hours per week × 52). The deliberate practice measure was log-transformed in Lehmann and Ericsson (1996) and Meinz and Hambrick (2010). No transformation was reported in the other studies, but there is evidence that log-transformations have only small effects on amount of variance explained (Macnamara, Hambrick, & Oswald, in preparation).

Five studies in Table 3 used piano sight-reading as a task, with various performance measures. Meinz's (2000) and Meinz and Hambrick's (2010) participants were pianists with a wide range of skill levels (beginning to advanced), and the measure of performance was the average of expert ratings of test performances across pieces. Our reliability estimates (.96 and .99, respectively) were the coefficient alphas for these measures. Tuffiash's (2002) participants were college undergraduates, including music majors and non-majors, and the performance measure was the highest level of difficulty at which a participant received a high rating (8/10 or higher) by both expert raters. Our reliability estimate (.80) was test–retest reliability for the first level of difficulty, which was based on the largest sample size. Lehmann and Ericsson's (1996) participants were highly skilled accompanists or piano soloists; an additional task was accompanying, and the performance measure for both tasks was the number of correctly played notes. Our reliability estimate (.88) was the correlation of number of correctly played notes with expert ratings of performance reported by Lehmann and Ericsson (1993). Kopiez and Lee's (2008) participants were current or past piano majors from a music institute, and the performance measure was again the number of correctly played notes. Reliability was not reported, and therefore our reliability estimate (.88) was the same as for Lehmann and Ericsson (1996).

Table 3
Correlations between deliberate practice and music performance.

Study	Sample	Performance measure	N	r (95% CI)	$\hat{r}$ (95% CI)	% var.
Lehmann and Ericsson (1996)	Accompanists and piano soloists	Number of correctly played notes	16	.36 (−.17, .73)	.43 (−.20, .87)	18.5 (81.5)
Meinz (2000)	Beginning to advanced pianists	Expert ratings of performances	107	.41 (.24, .56)	.47 (.27, .64)	22.1 (77.9)
Tuffiash (2002)	Undergraduate music and non-music majors	Expert ratings of performances	135	.58 (.46, .68)	.73 (.58, .85)	53.3 (46.7)
Kopiez and Lee (2008)	Music institute piano majors and graduates	Number of correctly performed pitches	52	.25 (−.02, .49)	.30 (−.02, .58)	9.0 (91.0)
Ruthsatz et al. (2008) – Study 1	High school band members	Rank based on band director rating of musical achievement	178	.34 (.20, .46)	.43 (.25, .58)	18.5 (81.5)
Ruthsatz et al. (2008) – Study 2A	Music institute students	Rank based on conservatory audition score	64	.31 (.07, .52)	.39 (.09, .65)	15.2 (84.8)
Ruthsatz et al. (2008) – Study 2B	University music majors	Rank based on faculty rating of musical achievement	19	.54 (.11, .80)	.68 (.14, 1.00)	46.2 (53.8)
Meinz and Hambrick (2010)	Beginning to advanced pianists	Expert ratings of performances	57	.67 (.50, .79)	.75 (.56, .89)	56.3 (43.7)

Note. Studies are in chronological order. r, correlation between cumulative deliberate practice and music performance. $\hat{r}$, corrected correlation. % var., percentage of variance in music performance explained (vs. unexplained) by deliberate practice after correction for measurement error variance ($\hat{r}^2$). Reliability for deliberate practice (r_{xx}), .80. Reliability for music performance (r_{yy}): Lehmann & Ericsson, 1996 (.88); Meinz, 2000 (.96); Tuffiash, 2002 (.80); Kopiez & Lee, 2008 (.88); Ruthsatz et al., 2008 (.80); Meinz & Hambrick, 2010 (.99).

The remaining three studies in Table 3–Study 1, Study 2A, and Study 2B of Ruthsatz, Detterman, Griscom, and Cirullo (2008)–reported global measures of musical performance. In Study 1, participants were high school band members, and the performance measure was rank in the band. In Study 2A, participants were music institute students, and the performance measure was rank based on audition score. In Study 2B, participants were university music majors and the performance measure was rank based on faculty rating of musical achievement. Our reliability estimate (.80) was based on previous demonstrations of internal consistency, test–retest, and inter-rater reliability of .80 or higher for similar measures of musical achievement (e.g., Bergee, 2003; Hash, 2012; Kinney, 2009; Lien & Humphreys, 2001). (Note that in the Ruthsatz et al. studies deliberate practice included practice and lessons.)

There is no standardized measure of performance in music as there is in chess, but there was an extremely wide range of music skill in some of the studies. Tuffiash's (2002) sample included both non-music majors recruited from an introductory psychology course and music majors from one of the top colleges of music in the U.S., and Meinz's (2000) and Meinz and Hambrick's (2010) participants ranged from beginner to professional.

4.1. Results of music reanalyses

Correlations between deliberate practice and music performance, with 95% confidence intervals (CIs), are displayed in Table 3. On average, deliberate practice explained 29.9% of the variance in performance after correcting for measurement error variance (avg. $\hat{r} = .52$; sample size-weighted avg. $\hat{r} = .52$), leaving 70.1% of the variance unexplained and potentially explainable by other factors (see Fig. 3). The 95% CI included 1.0 in Ruthsatz et al. (2008), Study 2B, but the confidence interval was extremely wide ($\hat{r} = .68$, 95% CI, .14, 1.0) due to a small sample size ($N = 19$). Note also that the correlation was much smaller in Ruthsatz et al.'s Study 2A ($\hat{r} = .39$, 95% CI, .09, .65), which used the same method but a larger sample ($N = 64$).

There was a wide age range in two studies, Meinz (2000) and Meinz and Hambrick (2010), but partialing age had almost no effect on the corrected correlations (change in $\hat{r}$, .02 and .00, respectively). The age ranges were much narrower in other studies, as the samples mainly included similar-aged students, and age correlations were not reported.

We again performed an analysis in which we assumed different levels of reliability for deliberate practice. Results are shown in Table 4. As for chess, even assuming a low level of reliability ($r_{xx} = .60$), deliberate practice left a very large proportion of the variance unexplained (60.4%).

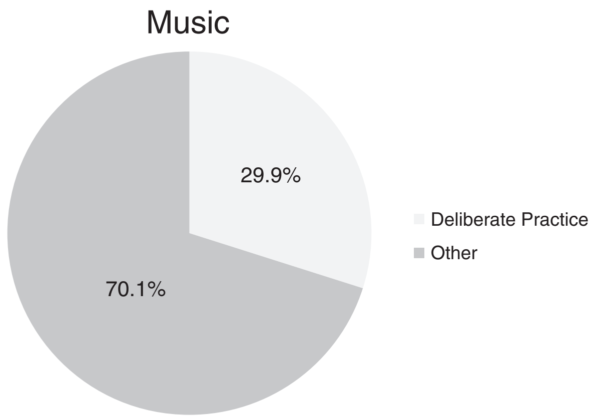

Fig. 3. Average percentage of variance in music performance accounted for by deliberate practice, correcting for measurement error variance. The light gray region represents reliable variance explained by deliberate practice; the dark gray region represents reliable variance not explained by deliberate practice.

4.2. Discussion of music reanalyses

On average across studies, deliberate practice explained about 30% of the reliable variance in music performance, leaving about 70% unexplained and potentially explainable by other factors. We conclude that deliberate practice is not sufficient to account for individual differences in music performance.

Results of other studies provide further support for this conclusion. Simonton (1991) found a large amount of variability in the amount of time it took famous classical composers to have their first "hit," and that the interval between the first composition and the first hit correlated significantly and negatively with maximum annual output, lifetime productivity, and posthumous reputation. Composers who rose to fame quickly–the most "talented"–had the most successful careers. Furthermore, Sloboda, Davidson, Howe, and Moore (1996) noted that although students at a selective music school ("high achievers") had accumulated more "formal practice" than students who were learning an instrument at a non-music school ("pleasure players"), there were some individuals at each skill level (grade) who did "less than 20 per cent of the mean amount of practice for that grade" and others who did "over four times as much practice than average to attain a given grade" (p. 301).

Ericsson et al.'s (1993) findings provide further support for our conclusion. Ericsson et al. did not report variability statistics for deliberate practice—no standard deviations, variances, or ranges.[7] However, the log-transformed values in their Fig. 15 indicate that deliberate practice in their study of pianists (Study 2) ranged from about 10,000 h to 30,000 h in the expert group.[8] The most practiced expert could have been no more than 11 years older than the least practiced expert (i.e., age 31 vs. 20), and yet the difference in deliberate practice between these subjects was about 20,000 h.[9] At 4 h a day,

[7] We contacted K. Anders Ericsson, Ralf Th. Krampe, and Clemens Tesch-Römer and requested their data; they are unable to provide it at this time.

[8] We digitized the three graphs in Ericsson et al.'s (1993) Fig. 15, extracted the values using Dagra's graphical extraction software, and back-transformed log accumulated hours of practice to accumulated hours of practice. With respect to Fig. 15, Ericsson et al. (1993) report very high correlations ($rs > -.85$) between cumulative deliberate practice and latency measures from a complex movement task—but these would have been highly inflated by the use of extreme-groups design in Study 2 (i.e., experts vs. amateurs).

[9] The age range for the expert pianists in Study 2, which Ericsson et al. (1993) noted was part of Ralf Krampe's dissertation, is not reported in Ericsson et al.'s article. We found this information in Krampe's dissertation and in Krampe and Ericsson's (1996) later report of data from this study; the age range is listed as 20 to 32 years in Krampe's dissertation and 20 to 31 years in Krampe and Ericsson (1996); we rely on the published report. We thank Daniela Regel, Max Planck Institute, for sending us a copy of Krampe (1994 Sigma edition).

Table 4
Results of reliability analysis for music performance.

	Reliability of deliberate practice (r_{xx})			
	$r_{xx} = .60$	$r_{xx} = .70$	$r_{xx} = .80$	$r_{xx} = .90$
Avg. % variance	39.6 (60.4)	33.8 (66.2)	29.9 (70.1)	26.3 (73.7)

Note. Avg. % variance, average percentage of variance in music performance explained (vs. unexplained) by deliberate practice.

365 days a year, it would take nearly 14 years to accumulate this amount of deliberate practice. It seems clear that some of Ericsson et al.'s pianists required much less deliberate practice than others to become experts.

5. General discussion

Deliberate practice does not explain all, nearly all, or even most of the variance in performance in chess and music, the two most widely studied domains in expertise research. Put another way, deliberate practice explains a considerable amount of the variance in performance in these domains, but leaves a much larger amount of the variance unexplained. The bottom line is that deliberate practice is necessary to account for why some people become experts in these domains and others fail to do so, but not even close to sufficient. So what else matters?

5.1. Starting age

Starting at a young age may be one factor. Ericsson et al. (1993) argued as follows that starting age influences performance insofar as it relates to the amount of deliberate practice an individual has accumulated: "An individual starting at an earlier age would have accumulated more deliberate practice and thus have acquired a higher level of performance" (p. 388). A testable prediction that follows from this statement is that the effect of starting age on performance should be mediated through deliberate practice. However, Gobet and Campitelli (2007) and Howard (2012) found that the effect of starting age on chess rating was not mediated through deliberate practice. That is, starting age correlated negatively with chess rating even after statistically controlling for deliberate practice, indicating that the players who started young tended to have an advantage as adult chess players independent of how much deliberate practice they had accumulated. Furthermore, in the study of composers mentioned earlier, Simonton (1991) found that compared with less eminent composers, the greatest composers started music lessons and composition lessons at a younger age and took less time to start making contributions to the repertoire. Taken together, this evidence suggests that there may be a critical period for acquiring complex skills just as there may be for acquiring language.

5.2. Intelligence

General intelligence and basic cognitive abilities–factors which are known to be highly heritable (Plomin, DeFries, McClearn, & McGuffin, 2008)–also appear to play a role in the acquisition of expert performance. One relevant construct is *working memory capacity*—the ability to maintain information in a highly active state (Engle, 2002). Meinz and Hambrick (2010) found that although deliberate practice accounted for nearly half (45.1%) of the variance in pianists' performance on a sight-reading task, working memory capacity accounted for an additional 7.4% of the variance—a statistically and practically significant effect. Ericsson and colleagues have argued that measures of working memory capacity themselves reflect acquired skills (Ericsson & Delaney, 1999; Ericsson & Kintsch, 1995), but working memory capacity and deliberate practice correlated near zero in this study ($r = .003$). There was also no evidence for a Deliberate Practice × Working Memory Capacity interaction, indicating that working memory capacity was no less important a predictor of performance for pianists with thousands of hours of deliberate practice than it was for beginners. In line with these findings, Kopiez and Lee (2006) reported that working memory capacity correlated significantly (one-tailed) with sight-reading performance at all but the most difficult level of music (Levels 1–4, $rs = .23$ to .32; Level 5, $r = .08$).[10] Thompson (1987) did not find a significant correlation between letter recall and sight-reading performance in flutists, but letter recall is more a test of short-term storage than working memory capacity (see, e.g., Engle, Tuholski, Laughlin, & Conway, 1999). Overall, it appears that working memory capacity has at least limited importance for musical expertise. An important question for future research is whether it plays a role in other musical skills (e.g., memorizing music).

Global measures of intelligence (IQ) have also been found to correlate with performance in chess and music, consistent with the possibility that a relatively high level of intelligence is necessary for success in these domains. Frydman and Lynn (1992) found that young chess players had an average performance IQ of 129, compared to a population average of 100, and that the average was higher for the best players (top-third avg. = 131) in the sample than the weakest players (bottom-third avg. = 124). Furthermore, Grabner, Neubauer, and Stern (2006) found that, even in highly rated players, IQ positively predicted performance on representative chess tasks (e.g., next best move). Bilalić et al. (2007) found that IQ was not a significant predictor of chess rating in the sample of elite young chess players listed in Table 1 after statistically controlling for practice. However, the sample size for the elite group was only 23, and mean IQ was significantly higher for the elite group ($M = 133$) than for the rest of the sample ($M = 114$). It has been suggested that chess training may transfer to IQ tests, but there is currently no compelling evidence for this (see Gobet & Campitelli, 2006, for a review). Instead, the effects of chess training appear to be domain-specific. For example, Schneider, Gruber, Gold, and Opwis (1993) found that children who played chess outperformed adults in a chessboard memory task, whereas the adults outperformed the children in a digit recall task.

IQ correlates positively with music performance, as well. Luce (1965) found a correlation of .53 ($p < .01$) between IQ and sight-reading performance in high school band members, and Salis (1977) reported a correlation of .58 between these variables in a university sample. Gromko (2004) found positive

[10] Kopiez and Lee (2006) reported a correlation of .258 between Level 4 performance and working memory, but did not flag this value as statistically significant in their Table 7 (p. 109). This is apparently a typo, because as they note, one-tailed $p = .032$ for this correlation, with $df = 50$.

correlations between both verbal ability and spatial ability ($rs = .35–.49$) and sight-reading performance in high school wind players, and Hayward and Gromko (2009) found a significant positive correlation ($r = .24$) between a measure of spatial ability based on three ETS tests and sight-reading performance in university wind players. Ruthsatz et al. (2008) found that Raven's scores correlated positively and significantly with musical achievement in high school band members ($r = .25$). This correlation was not statistically significant in a sample of more highly accomplished conservatory students and music majors, but this could have been due to a ceiling effect on Raven's, as these participants had been heavily selected for cognitive ability.

Ruthsatz and Detterman (2003) documented the case of a 6-year old piano prodigy named “Derek,” who at the time of the study had played in numerous concerts, appeared on national television, and released two CDs of his music. Derek scored at or above the 95th percentile on a test of musical aptitude, and had not engaged in *any* activity that would qualify as deliberate practice. Derek did, however, score well above the average on subsets of the Stanford–Binet Intelligence Scale: verbal reasoning (130), abstract reasoning (114), quantitative reasoning (120), and short-term memory (158). More recently, Ruthsatz and Urbach (2012) administered a standardized IQ test (the Stanford–Binet) to eight child prodigies, six of whom were musical prodigies. Despite full-scale IQs that ranged from 108 to 147–just above average to above the conventional cutoff for “genius”–all of the prodigies were at or above the 99th percentile for working memory (indeed, six scored at the 99.9th percentile).

The results of the landmark Study of Mathematically Precocious Youth (SMPY) are generally relevant to this discussion (see Robertson, Smeets, Lubinski, & Benbow, 2010). As part of a youth talent search, a large sample of children took the SAT by age 13, and those scoring in the top 1% were identified and tracked over the next two decades. Remarkably, individual differences in SAT scores–even within this highly restricted range of ability–predicted individual differences in scientific achievements. For example, compared to participants who were “only” in the 99.1 percentile for overall SAT score–which largely reflects general intelligence (Frey & Detterman, 2004)–those participants who had scored in the 99.9 percentile–the profoundly gifted–were 3.6 times more likely to have earned a doctorate, 5 times more likely to have published an article in a STEM journal, and 3 times more likely to have registered a patent (Lubinski, 2009).

General intelligence does not *always* predict performance. In a study of football players, Lyons, Hoffman, and Michel (2009) found that scores on the *Wonderlic Personnel Test*, a widely administered group intelligence test, correlated essentially zero with success in the National Football League, even in the quarterback position, which is believed to place the highest demand on information processing. Furthermore, Hambrick et al. (2012) found that spatial ability positively predicted success in a complex geological problem solving task in novice geologists, but not in experts. There is a clear need to develop theories of expert performance that take into account how the contribution of cognitive ability factors to performance varies across domains (e.g., cognitive vs. physical domains), and across situations or tasks in a given domain.

6. Personality

Ericsson et al. (1993) hypothesized as follows that personality factors may have an indirect effect on the acquisition of expert performance through deliberate practice:

> within our framework we would expect that several 'personality' factors, such as individual differences in activity levels and emotionality may differentially predispose individuals toward deliberate practice as well as allow these individuals to sustain very high levels of it for extended periods (p. 393).

There is now evidence to support this one part of the deliberate practice view. In a study of Spelling Bee contestants, Duckworth, Kirby, Tsukayama, Berstein, and Ericsson (2012) found that "grit"–a personality factor reflecting persistence in accomplishing long-term goals–positively predicted deliberate practice, which in turn positively predicted spelling performance. Similarly, in a study of classical musicians, Bonneville-Roussy, Lavigne, and Vallerand (2011) found that “passion” positively predicted “mastery goals,” which in turn positively predicted deliberate practice, which in turn positively predicted music performance.

This evidence suggests that personality is an important part of the expert performance puzzle. But, critically, this evidence does *not* suggest that personality is one of the “other” factors that accounts for the variance in performance that deliberate practice leaves unexplained (see the dark gray region of Figs. 1 and 3). That is, effects of the personality factors on performance in these studies were, as Ericsson (2012b) stressed in a discussion of the Duckworth et al. (2012) study, fully mediated through and thus completely explained by individual differences in deliberate practice. So, personality factors may explain why some people engage in more deliberate practice than others, but they do not appear to independently explain individual differences in performance.

6.1. Genes

There is some evidence that individual differences in performance are heritable. In the National Merit twin sample, Coon and Carey (1989) found heritability estimates of 38% for males and 20% for females for a measure of music achievement based on honors in music contests. Vinkhuyzen, van der Sluis, Posthuma, and Boomsma (2009) analyzed data from a study in which 1685 twin pairs rated their competence in chess, music, and several other domains on a scale from 1 (less competent than most people) to 4 (exceptionally skilled). For endorsement of “exceptional talent” (rating of 4 vs. 1, 2, or 3), heritability ranged from 50% to 92%.

There is also evidence for genetic effects on creativity. Based on correlations between scientific training/performance variables and both personality and intellectual traits, Simonton (2008) found nontrivial (lower-bound) heritability estimates for scientific training/performance variables—for example, about 24% for comprehensive exam scores and about 10% for faculty ratings. An important goal for future behavior genetic research on expert performance is to investigate whether there are genetic effects on objective measures of performance (e.g., music tasks, scientific problem solving tasks), and to investigate whether such effects on performance are accounted for by intelligence, personality, or both.

7. Popular myths about expert performance

Two myths regarding deliberate practice and expert performance have taken root in the public's imagination. The first myth is that people require very similar amounts of deliberate practice to acquire expert performance. Gladwell (2008) wrote in *Outliers* that Ericsson et al.'s (1993) "research suggests that once a musician has enough ability to get into a top music school, the thing that distinguishes one performer from another is how hard he or she works. That's it." (p. 39). Similarly, Syed (2010) wrote in *Bounce* that

> Top performers had devoted thousands of additional hours to the task of becoming master performers. But that's not all. Ericsson also found that there were no exceptions to this pattern: nobody who had reached the elite group without copious practice, and nobody who had worked their socks off but failed to excel. (pp. 12-13)

Such categorical claims are incorrect. The evidence is quite clear that some people *do* reach an elite level of performance without copious practice, while other people fail to do so *despite* copious practice.

The second myth is that it requires at least ten years, or 10,000 hours, of deliberate practice to reach an elite level of performance. Ericsson et al. (2007) explained this idea as follows: "Our research shows that even the most gifted performers need a *minimum* of ten years (or 10,000 hours) of intense training before they win international competitions" (p. 119, emphasis added). Subsequently, Gladwell (2008) proposed in *Outliers* that "Ten thousand hours is the magic number of greatness" (p. 41). More recently, the Nobel laureate Daniel Kahneman (2011) wrote in his book *Thinking, Fast and Slow* that "Studies of chess masters have shown that at least 10,000 hours of dedicated practice…are required to attain the highest levels of performance" (p. 238). But the data indicate that there is an *enormous* amount of variability in deliberate practice—even in elite performers. One player in Gobet and Campitelli's (2007) chess sample took 26 years of serious involvement in chess to reach a master level, while another player took less than *2 years* to reach this level.

Some normally functioning people may *never* acquire expert performance in certain domains, regardless of the amount of deliberate practice they accumulate. In Gobet and Campitelli's (2007) chess sample, four participants estimated more than 10,000 h of deliberate practice, and yet remained intermediate-level players. This conclusion runs counter to the egalitarian view that anyone can achieve most anything he or she wishes, with enough hard work. The silver lining, we believe, is that when people are given an accurate assessment of their abilities and of the likelihood of achieving certain goals given those abilities, they may gravitate towards domains in which they have a realistic chance of becoming an expert through deliberate practice.

8. Beyond the deliberate practice view

The debate over why and how some people become experts and others fail to do so has been a topic of intense debate in psychology for well over a century, and it will remain so for many years to come. The intensity of the debate to this point likely reflects a clash between what Cronbach (1957) called the two "disciplines" of scientific psychology—experimental psychology and differential psychology. Expertise researchers trained as experimental psychologists often seek to identify general principles that account for expert performance and to treat individual differences as "error," whereas those trained as differential psychologists seek to identify factors that account for individual differences. But as Simonton (1999) advises, "psychology must endeavor to identify all of the significant causal factors behind exceptional performance rather than merely rest content with whatever factor happens to account for the most variance" (p. 454). For researchers seriously interested in advancing the science of expert performance, the task now is to rise above disciplinary and ideological differences and develop and rigorously test theories that take into account as many potentially relevant explanatory factors as possible—including not only deliberate practice and other environmental variables, but heritable traits such as general intelligence, and task and situational factors that may moderate effects of individual-difference variables on performance.

An open-minded exchange of ideas among researchers with diverse theoretical and methodological perspectives will make this possible and shed fresh light on the origins of expert performance.

References

Ackerman, P. L. (2014). Nonsense, common sense, and science of expert performance: Talent and individual differences. *Intelligence*, *45*, 6–17.

Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.). New York: John Wiley and Sons.

Bergee, M. J. (2003). Faculty interjudge reliability of music performance evaluation. *Journal of Research in Music Education*, *51*, 137–150. http://dx.doi.org/10.2307/3345847.

Bilalić, M., McLeod, P., & Gobet, F. (2007). Does chess need intelligence? A study with young chess players. *Intelligence*, *35*, 457–470. http://dx.doi.org/10.1016/j.intell.2006.09.005.

Bonneville-Roussy, A., Lavigne, G. L., & Vallerand, R. J. (2011). When passion leads to excellence: The case of musicians. *Psychology of Music*, *39*, 123–138. http://dx.doi.org/10.1177/0305735609352441.

Brooks, D. (2011). *The social animal: The hidden sources of love, character, and achievement*. New York: Random House.

Burns, B. (2004). The effects of speed on skilled chess performance. *Psychological Science*, *15*, 442–447. http://dx.doi.org/10.1111/j.0956-7976.2004.00699.x.

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.

Charness, N., Krampe, R. Th., & 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. http://dx.doi.org/10.1002/acp.1106.

Colvin, G. (2010). *Talent is overrated. What really separates world-class performers from everybody else*. New York: Penguin.

Coon, H., & Carey, G. (1989). Genetic and environmental determinants of musical ability in twins. *Behavior Genetics*, *19*, 183–193. http://dx.doi.org/10.1007/BF01065903.

Coyle, D. (2009). *The talent code: It's born. It's grown. Here's how.* New York: Bantam.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684. http://dx.doi.org/10.1037/h0043943.

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: HarperCollins.

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.

de Bruin, A. B. H., Smits, N., Rikers, R. M. J. P., & Schmidt, H. G. (2008). Deliberate practice predicts performance over time in adolescent chess players and drop-outs: A linear mixed model analysis. *British Journal of Psychology*, *99*, 473–497. http://dx.doi.org/10.1348/000712608X295631.

Detterman, D. K., Gabriel, L. T., & Ruthsatz, J. M. (1998). Absurd environmentalism. *The Behavioral and Brain Sciences*, *21*, 411–412. http://dx.doi.org/10.1017/S0140525X98271238.

Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2012). Deliberate practice spells success: Why grittier competitors triumph at the National Spelling Bee. *Social Psychological and Personality Science*, *2*, 174–181. http://dx.doi.org/10.1177/1948550610385872.

Elo, A. E. (1986). *The rating of chessplayers, past and present* (2nd ed.). New York: Arco.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. http://dx.doi.org/10.1111/1467-8721.00160.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology. General*, *128*, 309–331.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683–703). New York: Cambridge University Press.

Ericsson, K. A. (2007). Deliberate practice and the modifiability of body and mind: Toward a science of the structure and acquisition of expert and elite performance. *International Journal of Sport Psychology*, *38*, 4–34.

Ericsson, K. A. (2012a). Training history, deliberate practice and elite sports performance: an analysis in response to Tucker and Collins review–what makes champions? *British Journal of Sports Medicine*, *47*, 533–535.

Ericsson, K. A. (2012b). *The danger of delegating education to journalists: Why the APS Observer needs peer review when summarizing new scientific developments.* Unpublished manuscript. Retrieved from http://www.psy.fsu.edu/faculty/ericsson/ericsson.hp.html

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

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake, & P. Shah (Eds.), *Models of working memory* (pp. 257–297). Cambridge, MA: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245. http://dx.doi.org/10.1037/0033-295X.102.2.211.

Ericsson, K. A., Krampe, R. Th., & 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.

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. http://dx.doi.org/10.4219/jeg-2005-335.

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review*, *85*, 114–121.

Freeman, J. (2007). If you can't measure it—it doesn't exist. *High Ability Studies*, *18*, 65–66. http://dx.doi.org/10.1080/13598130701350726.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*, 373–378. http://dx.doi.org/10.1111/j.0956-7976.2004.00687.x.

Frydman, M., & Lynn, R. (1992). The general intelligence and spatial abilities of gifted young Belgian chess players. *British Journal of Psychology*, *83*, 233–235. http://dx.doi.org/10.1111/j.2044-8295.1992.tb02437.x.

Galton, F. (1869). *Hereditary genius.* London: Macmillan.

Gardner, H. (1995). "Expert performance: Its structure and acquisition": Comment. *American Psychologist*, *50*, 802–803. http://dx.doi.org/10.1037/0003-066X.50.9.802.

Gladwell, M. (2008). *Outliers: The story of success.* New York: Little, Brown, and Co.

Gobet, F., & Campitelli, G. (2006). Educational benefits of chess instruction: A critical review. In T. Redman (Ed.), *Chess and education: Selected essays from the Koltanowski conference* (pp. 81–97). Dallas, TX: Chess Program at the University of Texas at Dallas.

Gobet, F., & Campitelli, G. (2007). The role of domain-specific practice, handedness, and starting age in chess. *Developmental Psychology*, *43*, 159–172. http://dx.doi.org/10.1037/0012-1649.43.1.159.

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. http://dx.doi.org/10.1016/j.brainresbull.2006.02.009.

Grabner, R. H., Stern, E., & Neubauer, A. C. (2007). Individual differences in chess expertise: A psychometric investigation. *Acta Psychologica*, *124*, 398–420. http://dx.doi.org/10.1016/j.actpsy.2006.07.008.

Gromko, J. E. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education*, *52*, 6–15. http://dx.doi.org/10.2307/3345521.

Hambrick, D. Z., Libarkin, J. C., Petcovic, H. L., Baker, K. M., Elkins, J., Callahan, C. N., Turner, S. P., Rench, T. A., & LaDue, N. D. (2012). A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. *Journal of Experimental Psychology. General*, *141*, 397–403. http://dx.doi.org/10.1037/a0025927.

Hambrick, D. Z., & Meinz, E. J. (2011a). Sorry, strivers. Talent matters. *The New York Times*. Sunday Review, 12.

Hambrick, D. Z., & Meinz, E. J. (2011b). Limits of the predictive power of domain-specific experience. *Current Directions in Psychological Science*, *20*, 275–279. http://dx.doi.org/10.1177/0963721411422061.

Hambrick, D. Z., & Meinz, E. J. (2012). Working memory capacity and musical skill. In T. P. Alloway, & R. G. Alloway (Eds.), *Working memory: The connected intelligence* (pp. 137–155). New York: Psychology Press.

Hash, P. M. (2012). An analysis of the ratings and interrater reliability of high school band contests. *Journal of Research in Music Education*, *60*, 81–100. http://dx.doi.org/10.1177/0022429411434932.

Hayward, C. M., & Gromko, J. E. (2009). Relationships among music sight-reading and technical proficiency, spatial visualization, and aural discrimination. *Journal of Research in Music Education*, *57*, 29–36. http://dx.doi.org/10.1177/0022429409332677.

Howard, R. W. (2011). Does high-level intellectual performance depend on practice alone? Debunking the Polgár sisters case. *Cognitive Development*, *26*, 196–202. http://dx.doi.org/10.1016/j.cogdev.2011.04.001.

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.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park: Sage.

Kahneman, D. (2011). *Thinking, fast and slow.* New York: Farrar, Straus, and Giroux.

Kaufman, S. B. (2007). Investigating the role of domain general mechanisms in the acquisition of domain specific expertise. *High Ability Studies*, *18*, 71–73. http://dx.doi.org/10.1080/13598130701350767.

Kinney, D. W. (2009). Internal consistency of performance evaluations as a function of music expertise and excerpt familiarity. *Journal of Research in Music Education*, *56*, 322–337. http://dx.doi.org/10.1177/0022429408328934.

Kopiez, R., & Lee, J. I. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research*, *8*, 97–120. http://dx.doi.org/10.1080/14613800600570785.

Kopiez, R., & Lee, J. I. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research*, *10*, 41–62. http://dx.doi.org/10.1080/14613800701871363.

Krampe, R. Th. (1994). *Maintaining excellence: Cognitive-motor performance in pianists differing in age and skill level.* Berlin, Germany: Max-Planck-Institut für Bildungsforschung.

Krampe, R. Th., & 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.

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology: A Journal of Research in Music Cognition*, *12*, 182–195. http://dx.doi.org/10.1037/h0094108.

Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychmusicology: A Journal of Research in Music Cognition*, *15*, 1–29. http://dx.doi.org/10.1037/h0094082.

Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession.* New York: Penguin.

Lien, J. L., & Humphreys, J. T. (2001). Relations among selected variables in the South Dakota all-state band auditions. *Journal of Research in Music Education*, *49*, 146–155. http://dx.doi.org/10.2307/3345866.

Lubinski, D. (2009). Exceptional cognitive ability: The phenotype. *Behavior Genetics*, *39*, 350–358. http://dx.doi.org/10.1007/s10519-009-9273-0.

Luce, J. R. (1965). Sight-reading and ear-playing abilities as related to instrumental music students. *Journal of Research in Music Education*, *13*, 101–109. http://dx.doi.org/10.2307/3344447.

Lyons, B., Hoffman, B., & Michel, J. (2009). Not much more than g? An examination of the impact of intelligence on NFL performance. *Human Performance*, *22*, 225–245. http://dx.doi.org/10.1080/08959280902970401.

Macnamara, B., Hambrick, D. Z., & Oswald, F. L. (2013). *Deliberate practice and performance in music, sports, games, professions, and education: A meta-analysis.* (in preparation).

Marcus, G. (2012). *Guitar zero: The science of becoming musical at any age.* New York: Penguin.

Meinz, E. J. (2000). Experience-based attenuation of age-related differences in music cognition tasks. *Psychology and Aging*, *15*, 297–312. http://dx.doi.org/10.1037/0882-7974.15.2.297.

Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological Science*, *21*, 914–919. http://dx.doi.org/10.1177/0956797610373933.

Pink, D. H. (2009). *Drive. The surprising truth about what motivates us.* New York: Riverhead Books.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral genetics* (5th ed.). New York: Worth Publishers.

Robertson, K. F., Smeets, S., Lubinski, D., & Benbow, C. P. (2010). Beyond the threshold hypothesis: Even among gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence. *Current Directions in Psychological Science*, *19*, 346–351. http://dx.doi.org/10.1177/0963721410391442.

Ruthsatz, J., & Detterman, D. K. (2003). An extraordinary memory: The case study of a musical prodigy. *Intelligence*, *31*, 509–518. http://dx.doi.org/10.1016/S0160-2896(03)00050-3.

Ruthsatz, J., Detterman, D. K., Griscom, W. S., & Cirullo, B. A. (2008). Becoming an expert in the musical domain: It takes more than just practice. *Intelligence*, *36*, 330–338. http://dx.doi.org/10.1016/j.intell.2007.08.003.

Ruthsatz, J., & Urbach, J. B. (2012). Child prodigy: A novel cognitive profile places elevated general intelligence, exceptional working memory and attention to detail at the root of prodigiousness. *Intelligence*, *40*, 419–426. http://dx.doi.org/10.1016/j.intell.2012.06.002.

Salis, D. L. (1977). *The identification and assessment of cognitive variables associated with reading of advanced music at the piano (unpublished doctoral dissertation).* Pittsburgh, PA: University of Pittsburgh.

Schmidt, F., & Hunter, J. (1999). Theory testing and measurement error. *Intelligence*, *27*, 183–198. http://dx.doi.org/10.1016/S0160-2896(99)00024-0.

Schneider, W. (1998). Innate talent or deliberate practice as determinants of exceptional performance: Are we asking the right question? *The Behavioral and Brain Sciences*, *21*, 423–424. http://dx.doi.org/10.1017/S0140525X98411233.

Schneider, W., Gruber, H., Gold, A., & Opwis, K. (1993). Chess expertise and memory for chess positions in children and adults. *Journal of Experimental Child Psychology*, *56*, 328–349.

Shenk, D. (2010). *The genius in all of us. Why everything you've been told about genetics, talent, and IQ is wrong.* New York: Doubleday.

Simonton, D. K. (1991). Emergence and realization of genius: The lives and works of 120 classical composers. *Journal of Personality and Social Psychology*, *61*, 829–840. http://dx.doi.org/10.1037/0022-3514.61.5.829.

Simonton, D. K. (1999). Talent and its development: An emergenic and epigenetic model. *Psychological Review*, *106*, 435–457. http://dx.doi.org/10.1037/0033-295X.106.3.435.

Simonton, D. K. (2008). Scientific talent, training, and performance: Intellect, personality, and genetic endowment. *Review of General Psychology*, *12*, 28–46. http://dx.doi.org/10.1037/1089-2680.12.1.28.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology*, *87*, 287–309.

Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, *15*, 201–292. http://dx.doi.org/10.2307/1412107.

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: Lawrence Erlbaum.

Syed, M. (2010). *Bounce: Mozart, Federer, Picasso, Beckham, and the science of success.* New York: HarperCollins.

Thompson, W. B. (1987). Music sight-reading skill in flute players. *Journal of General Psychology*, *114*, 345–352. http://dx.doi.org/10.1080/00221309.1987.9711072.

Tuffiash, M. (2002). *Predicting individual differences in piano sight-reading skill: Practice, performance, and instruction (unpublished master's thesis).* Tallahassee, FL: Florida State University.

Tuffiash, M., Roring, R. W., & Ericsson, K. A. (2007). Expert performance in SCRABBLE: Implications for the study of the structure and acquisition of complex skills. *Journal of Experimental Psychology: Applied*, *13*, 124–134. http://dx.doi.org/10.1037/1076-898X.13.3.124.

van der Maas, H. L., & Wagenmakers, E. J. (2005). A psychometric analysis of chess expertise. *The American Journal of Psychology*, *118*, 29–60.

Vinkhuyzen, A. E., van der Sluis, S., Posthuma, D., & Boomsma, D. I. (2009). The heritability of aptitude and exceptional talent across different domains in adolescents and young adults. *Behavior Genetics*, *39*, 380–392. http://dx.doi.org/10.1007/s10519-009-9260-5.

Watson, J. B. (1930). *Behaviorism.* Chicago, IL: The University of Chicago Press.

Winner, E. (2000). The origins and ends of giftedness. *American Psychologist*, *55*, 159–169. http://dx.doi.org/10.1037/0003-066X.55.1.159.