The Impact of Domain-Specific Experience on Chess Skill: Reanalysis of a Key Study

ALEXANDER P. BURGOYNE and CHRISTOPHER D. NYE
Michigan State University

BROOKE N. MACNAMARA
Case Western Reserve University

NEIL CHARNESS
Florida State University

DAVID Z. HAMBRICK
Michigan State University

How important are training and other forms of domain-relevant experience in predicting individual differences in expertise? To answer this question, we used structural equation modeling to reanalyze data from a study of chess by Charness, Tuffiash, Krampe, Reingold, and Vasyukova (2005). Latent variables reflecting serious chess activity and formal instruction, along with a manifest variable indexing serious starting age, accounted for 63% of the variance in peak rating. Serious starting age had a significant negative effect on peak rating ($\beta = -.15$), even after we controlled for domain-specific experience, indicating an advantage for starting earlier. We also tested the prediction that formal instruction increases the effectiveness of serious study (Ericsson & Charness, 1994) using moderated regression. This claim was not supported. Overall, the results affirm that serious study and other forms of domain-specific experience are important pieces of the expertise puzzle, but other factors must matter too.

Supplemental materials are available at https://www.press.uillinois.edu/journals/ajp/media/chess_skill

KEYWORDS: chess, expertise, deliberate practice, starting age, formal instruction

How important are training and other forms of experience in accounting for individual differences in skill in complex domains? This question is the subject of vigorous debate in the scientific literature on expertise. A recent series of studies used meta-analysis to investigate this question. In the first meta-analysis, previous studies from two of the most widely researched domains of expertise, chess and music, were reanalyzed. After correction for measurement error (i.e., unreliability), estimates of deliberate practice (i.e., activities specifically designed to improve the current level of performance; Ericsson, Krampe, & Tesch-Römer, 1993) accounted for 34% of the variance in chess expertise and 30% of the variance in music expertise (Hambrick et al., 2014).

Subsequently, a second meta-analysis was conducted to estimate how much of the variance in skill could be accounted for by activities interpretable as deliberate practice in the domains of music, games, sports, education, and other professions. The results revealed that deliberate practice accounted for a sizable portion of variance in performance in games (26%), music (21%), and sports (18%) and for a small portion of variance in education (4%) and other professions (less than 1%) but in all domains left a larger proportion of variance unexplained (Macnamara, Hambrick, & Oswald, 2014). This was true even after liberal corrections for measurement error were applied to the average correlations.

PRESENT STUDY

As is often noted, one limitation of meta-analyses is that weaknesses in the design of individual studies can influence the overall results (Flather, Farkouh, Pogue, & Yusuf, 1997). With this in mind, the present study took a different approach and reanalyzed the results of a pair of studies of chess by Charness, Tuffiash, Krampe, Reingold, and Vasyukova (2005). These studies were exemplary in several respects. The sample sizes were very large for a study of expertise (*N*s = 239 and 180), expertise was determined by an objective measure of performance (i.e., Elo rating[1]; Elo, 1978), and the participants represented a wide range of chess skill (Elo rating range = 1150–2650). Furthermore, as Ericsson and Moxley (2012) emphasized, Charness et al. modeled their interview procedure for eliciting the estimates of serious study directly after the procedure used by Ericsson et al. (1993) to elicit estimates of deliberate practice in music. Overall, Ericsson (2005) described the evidence from this study as "the most compelling and detailed evidence for how designed training (deliberate practice) is the crucial factor in developing expert chess performance" (p. 237). Finally, this study collected estimates of multiple types of chess-specific activity, including not only serious study (the activity that most closely matches deliberate practice) but also tournament play and formal instruction.

Charness et al. (2005) reported three main findings, the first two of which were based on multiple regression analyses. The first finding was that total hours of serious study was a stronger predictor of current and peak chess rating than both formal instruction and tournament competition. The second finding was that predictors of current chess rating differed for younger (age less than 40) and older (age 40 or older) chess players. Many training activities significantly predicted younger chess players' current ratings, including total hours of serious study, total hours of tournament play, total years of private instruction, and current hours per week of serious study. By contrast, only total hours of serious study significantly predicted older chess players' current ratings. The third finding was that players who attained higher chess ratings engaged in more hours of serious study during the first decade of serious chess play.

In the present study, we used exploratory factor analysis (EFA) in conjunction with structural equation modeling (SEM) (Fabrigar, Wegener, MacCallum, & Strahan, 1999) to investigate the factor structure of the experience measures and then to assess the relative contributions of latent variables reflecting different forms of chess experience to chess skill (i.e., rating). A major advantage of SEM over other approaches to analyzing individual difference data (e.g., correlations, regression) is that it permits the researcher to analyze the data at the level of latent variables, which capture variance common to multiple measures of a construct and are free of random measurement error (Kline, 2011). Our major question of interest was how much of the variance in chess expertise could be explained by the different forms of chess experience measured in the Charness et al. (2005) project. The proportion of variance left unexplained indicates the extent to which unmeasured factors, such as cognitive ability (Burgoyne et al., 2016), can contribute to chess skill.

We addressed four questions in this study that were not addressed by Charness et al. (2005). The first question was whether the age at which people start seriously playing chess predicts skill, above and beyond chess experience factors. The existence of a sensitive (or critical) period has been postulated for language development (Lenneberg, 1967), normal vision (Hensch, 2004), second language acquisition (Johnson & Newport, 1989), and chess skill (Gobet & Campitelli, 2007), among other domains. Arpad

Elo, inventor of the Elo rating system for chess skill, suggested that early exposure to formal instruction and competition is critical to the attainment of chess mastery (Elo, 1978). Consistent with this hypothesis, Gobet and Campitelli (2007) found that the probability of becoming an international-level player (grandmaster or international master) was about 1 in 4 for players who started to play chess seriously before age 13 and only 1 in 55 for players who started at age 13 or older. Moreover, Gobet and Campitelli found a significant partial correlation between serious starting age and national rating after controlling for total practice hours ($r = -.40$, $p < .001$), indicating an advantage for starting earlier. In the present study, to further test the sensitive period hypothesis for chess, we used SEM to determine whether serious starting age predicts chess skill, after controlling for latent variables capturing different forms of domain-specific experience. If serious starting age predicts chess skill when experience is controlled for, this suggests that the advantage associated with starting earlier is not simply a result of having more time to train, as suggested by Ericsson et al. (1993), but that other factors (e.g., neural plasticity) may also play a role.

The second additional question was whether tournament play contributes to expertise above and beyond serious study and formal instruction. Tournament play (i.e., competition) is a form of domain-specific experience distinct from deliberate practice (see Ericsson et al., 1993). If tournament play independently contributes to chess rating, this would suggest that chess training programs should continue to augment serious study and formal instruction with other types of experience. To answer this question, we conducted a hierarchical regression analysis testing whether tournament play significantly predicted chess rating after taking into account serious study and formal instruction.

The third additional question was whether formal chess instruction moderates the effect of serious study on chess skill. Ericsson and colleagues argued that formal instruction increases the effect of practice on performance. For example, instructors can prescribe specific practice activities that are particularly beneficial for students. As Ericsson, Prietula, and Cokeley (2007) stated, "Having expert coaches makes a difference in a variety of ways. To start with, they can help you accelerate your learning process" (p. 120). Along the same lines, Ericsson and Charness (1994) speculated that serious study in chess is more effective when directed by a teacher:

> The activity of planning and extended evaluation of chess games is likely to improve a player's ability to internally represent chess positions, a memory skill that we discussed earlier in this article. This form of self-directed study has most of the characteristics of deliberate practice, but it is probably not as effective as individualized study guided by a skilled teacher. (p. 739)

This speculation leads to the prediction of an interaction between formal instruction and serious study on chess rating: There should be a greater increase in Elo rating per unit increase of serious study for chess players who have received formal instruction than for those who receive little or no formal instruction (Figure 1). We tested this prediction via moderated multiple regression.

The fourth additional question was whether predictors of peak chess rating differ for older and youn-

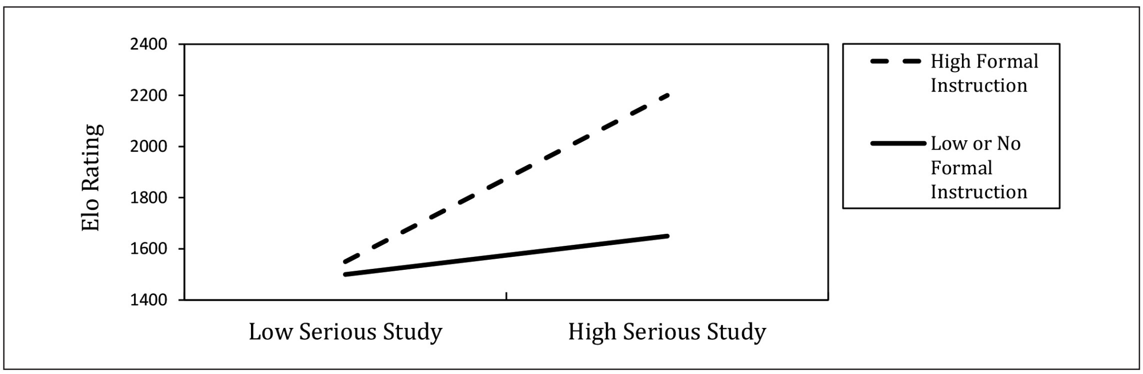

FIGURE 1. Hypothesized interaction between serious study and formal instruction based on Ericsson and Charness (1994)

ger chess players. Roring and Charness (2007) found that the average age at which players attained their peak rating was approximately 44 years old. Many of the participants included in this reanalysis are younger than 44 years old, and it is likely that some of them had not yet attained their true peak rating at the time of data collection. Moreover, Charness et al. (2005) found that predictors of *current* chess rating differed for older and younger players, with more factors contributing to younger chess players' current ratings than older chess players' current ratings. To answer this question, we performed regression analyses testing for predictors of peak rating separately for older (age 44 or older) and younger (age less than 44) chess players.

METHOD

Participants

To reiterate, the data for this study were from Study 1 and Study 2 of Charness et al. (2005). Study 1 consisted of 229 chess players (about 98% male) representing a wide range of skill (Elo rating M = 2046, SD = 290) and age (M = 36 years, SD = 15 years). A portion of this sample (n = 136) was previously analyzed in Charness, Krampe, and Mayr (1996). Study 2 consisted of 180 chess players (about 85% male) with an Elo rating of at least 1600 (M = 2009, SD = 254), who were at least 18 years old (M = 45 years, SD = 16 years). Charness et al. used stratified sampling in Study 2, so that participants' age and skill level were uncorrelated, and made sure that no participants from Study 1 were included in Study 2.

Materials

We obtained the raw deidentified data from Charness et al. (2005), which included a number of variables associated with players' time seriously studying, playing in tournaments, and receiving instruction and the number of chess books owned (assumed to be used for self-study), along with chess ratings and age information. See Table 1 for a brief description of each variable that we included in the present reanalysis; see Charness et al. (2005) for more information about the survey used to collect data.

Data Preparation

First, we combined the data from Studies 1 and 2 and used listwise deletion to remove any participants who were missing values for peak rating. We removed these participants because peak rating was our primary dependent measure of interest; 32 of the 409 participants were missing values for peak rating. Next, we used an expectation maximization algorithm (Dempster, Laird, & Rubin, 1977) to estimate the 129 remaining missing values (2.85% of the data) associated with other variables.[2]

RESULTS

Table 2 displays the descriptive statistics and correlations for all variables for the combined samples.

TABLE 1. Variable Names and Descriptions

Variable	Description
Peak rating	Highest attained Elo rating at time of data collection
Current rating	Elo rating at time of current age
Peak study time	Log hours of serious study up to age at peak rating
Total study time	Total log hours of serious study at current age
Peak tournament play	Log hours of tournament play up to age at peak rating
Total tournament play	Total log hours of tournament play at current age
Peak private instruction	Log years of private instruction up to age at peak rating
Total private instruction	Total log years of private instruction at current age
Peak group instruction	Log years of group instruction up to age at peak rating
Total group instruction	Total log years of group instruction at current age
Books	Log number of chess books owned at current age
Serious age	Age of beginning serious involvement in chess

TABLE 2. Correlations and Descriptive Statistics for Chess Skill, Activities, and Age Variables (Listwise $N = 377$)

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Peak rating		**.97**	**.58**	**.57**	**.43**	**.40**	**.38**	**.38**	**.30**	**.29**	**.58**	**−.39**
2. Current rating			**.57**	**.53**	**.43**	**.35**	**.39**	**.39**	**.28**	**.27**	**.56**	**−.40**
3. Peak study time				**.91**	**.60**	**.52**	**.25**	**.24**	**.17**	**.16**	**.47**	**−.17**
4. Total study time					**.54**	**.59**	**.25**	**.26**	**.18**	**.18**	**.49**	−.11
5. Peak tournament play						**.86**	**.20**	**.20**	**.19**	**.18**	**.33**	−.11
6. Total tournament play							**.18**	**.19**	**.21**	**.21**	**.34**	−.03
7. Peak private instruction								**1.00**	**.33**	**.33**	**.24**	**−.27**
8. Total private instruction									**.33**	**.33**	**.25**	**−.26**
9. Peak group instruction										**1.00**	**.26**	**−.34**
10. Total group instruction											**.26**	**−.34**
11. Books												**−.22**
12. Serious age												
Mean	2087	2038	3.41	3.52	3.45	3.55	0.18	0.19	0.33	0.34	1.74	15.9
SD	267	275	0.51	0.51	0.47	0.48	0.34	0.34	0.39	0.39	0.52	8.6

Note. *Coefficients in bold are statistically significant at* $p < .01$.

The training variables that had the highest correlations with peak rating were peak serious study time ($r = .58$, $p < .001$) and books ($r = .58$, $p < .001$). Serious age was negatively correlated with peak rating ($r = –.39$, $p < .001$), indicating that players who started playing chess seriously at an early age tended to reach a higher level of skill than players who started at a later age. Peak rating and current rating were highly correlated ($r = .97$, $p < .001$) and yielded nearly identical results in all analyses (for the results of the current rating analysis and for histograms depicting the distribution of variables used in the main analyses, see the Supplemental Materials available online).

Exploratory Factor Analyses

We used EFA to investigate the factor structure of predictors of peak rating. Predictors of peak rating included peak study time, peak tournament time, books, peak private instruction, and peak group instruction. We performed the EFA using principal axis extraction. The criterion for factor extraction was an eigenvalue greater than 1.0, and we rotated the factors with an oblique rotation procedure (Promax) to allow any factors that emerged to correlate.

Results are presented in Table 3. The analysis yielded two factors. The first factor, which we label serious chess activity, had high loadings on the following variables: peak serious study time, peak tournament time, and books. The second factor, which we label chess instruction, had high loadings on the following variables: peak private instruction and peak group instruction. The correlation between the two factors was $r = .40$.

TABLE 3. Summary of Exploratory Factor Analysis Results for Predictors of Peak Rating (Listwise $N = 377$)

Variable	Serious chess activity	Chess instruction
Peak study time	**.98**	–.09
Peak tournament play	**.62**	.05
Books	**.42**	.24
Peak private instruction	.11	**.46**
Peak group instruction	–.04	**.68**
Eigenvalues	2.25	1.05
Percentage of variance	44.98	20.97

Note. *Factor loadings ≥.40 appear in bold.*

Structural Equation Modeling

To reiterate, the primary focus of this study was to assess how much of the variance in peak rating could be accounted for by domain-specific experience, including deliberate practice activities. We used SEM to address this goal. Two steps were involved. Guided by the results of the EFA, the first step was to perform confirmatory factor analysis (CFA) on predictors of peak rating to assess the fit of the two-factor model to the data. We specified two factors in the CFA: serious chess activity and chess instruction (each indicator loaded on the latent factor identified in the EFA). Model fit was good, $\chi^2(4) = 15.78$, $p = .003$, confirmatory fit index = .97, normed fit index = .96, RMSEA = .09.

The second step in the SEM was to assess the effect of serious chess activity and chess instruction, along with serious starting age, on peak rating. Results are illustrated in Figure 2. Serious chess activity had a significant positive effect on rating ($\beta = .62$, $p < .001$), whereas the effect of chess instruction was small ($\beta = .19$, $p = .035$). Furthermore, serious starting age had a significant negative effect on rating ($\beta = -.15$, $p = .005$), above and beyond the chess experience factors. Collectively, the model accounted for 63.4% of the variance in peak rating. Model fit was acceptable, $\chi^2(10) = 51.82$, $p < .001$, confirmatory fit index = .94, normed fit index = .93, RMSEA = .11.

Additional Analyses

COMPETITION EXPERIENCE.

We used hierarchical multiple regression to examine whether competition experience (i.e., peak tournament play) contributed to the prediction of peak rating, above and beyond serious study and formal instruction. We entered peak serious study time, peak private instruction, and peak group instruction in Step 1 of the model and peak tournament play in Step 2 of the model.

The overall model accounted for 42.1% of the variance in peak rating, $F(4, 372) = 67.66$, standard error of estimate (SEE) = 204, $p < .001$ (Table 4). Although the effect size was small, tournament play contributed significantly to the prediction of peak rating, above and beyond study time and formal instruction, $\Delta R^2 = .006$, $p = .049$.[3] Higher levels of tournament play were associated with higher peak rating.

SERIOUS STUDY AND FORMAL INSTRUCTION.

We used multiple regression to test the hypothesis that the effect of serious study on peak rating was moderated by formal instruction such as private lessons or group lessons. We also tested whether the effect of serious study on peak rating was moderated by chess books, another training aid. We took standardized scores for peak serious study time, peak private instruction, peak group instruction, and books,

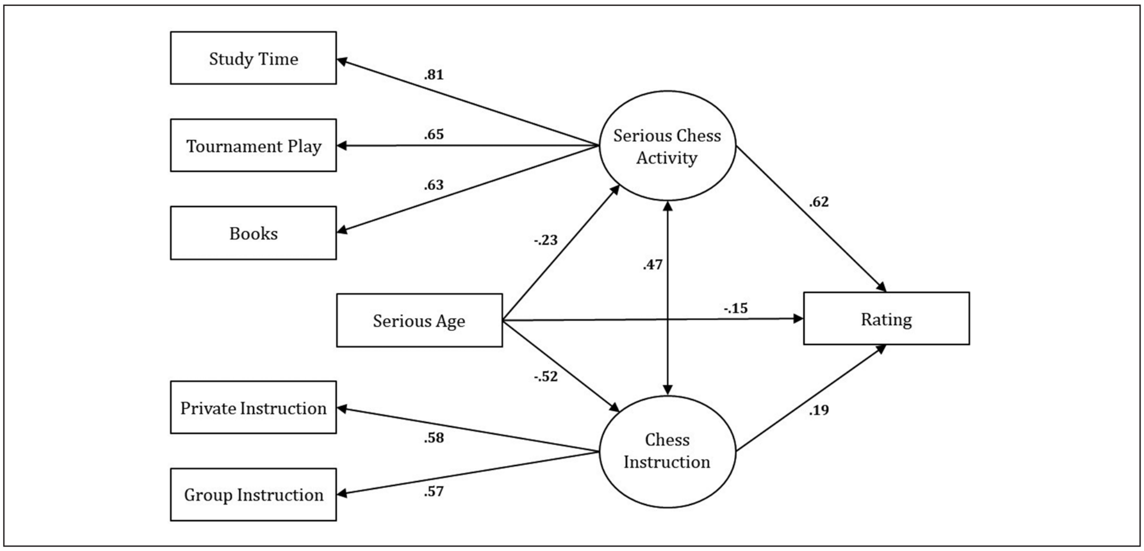

FIGURE 2. Structural equation model predicting peak rating ($N = 377$)

TABLE 4. Hierarchical regression of peak rating on serious study, formal instruction, and tournament play for tournament-rated chess players ($N = 377$)

	ΔR^2	Sig. *F* Change	*B*	β	*t*	*p*
Step 1	.415	< .001				
Peak study time			234.0	.45	9.02	< .001
Peak private instruction			159.7	.20	4.71	< .001
Peak group instruction			92.1	.13	3.17	.002
Step 2	.006	< .05				
Peak tournament play			55.5	.10	1.97	.049

Note. B, *unstandardized regression coefficient;* β, *standardized regression coefficient.* Bs, βs, ts, *and* ps *are reported for the full model.*

and computed peak serious study time × peak private instruction, peak serious study time × peak group instruction, and peak serious study time × books interaction terms (Aiken, West, & Reno, 1991). Next, we entered peak serious study time, peak tournament play, peak private instruction, peak group instruction, books, and serious starting age in Step 1 of the model and the preceding interaction terms in Step 2.

The overall model accounted for 54.8% of the variance in peak rating, $F(9, 367) = 49.53$, SEE = 181, $p < .001$ (Table 5). The interaction terms did not contribute significantly to the prediction of peak rating, $\Delta R^2 = .002$, $p = .67$.

AGE SUBSETS.

We conducted a series of multiple regression analyses on predictors of peak rating for older (age 44 or older) and younger (age less than 44) chess players (Table 6). For older chess players, the model accounted for 54.0% of the variance in peak rating, $F(6, 95) = 18.59$, SEE = 178, $p < .001$. For younger chess players, the model accounted for 54.3% of the variance in peak rating, $F(6, 219) = 43.41$, SEE = 180, $p < .001$. For younger chess players, serious study time, tournament play, private instruction, books, and serious age were significant predictors of peak rating (all positive, except for serious age). For older chess players, only

TABLE 5. Hierarchical regression of peak rating on predictor variables and interaction terms ($N = 377$)

	ΔR^2	Sig. *F* Change	*B*	β	*t*	*p*
Step 1	.547	< .001				
Peak study time			154.4	.30	6.25	< .001
Peak tournament play			47.8	.08	1.90	.058
Peak private instruction			126.3	.16	3.56	< .001
Peak group instruction			15.2	.02	0.55	.581
Books			164.2	.32	7.83	< .001
Serious age			–6.7	–.22	–5.55	< .001
Step 2	.002	.67				
Study × Private Instruction			–10.9	–.04	–0.95	.341
Study × Group Instruction			9.2	.04	0.87	.386
Study × Books			–4.1	–.02	–0.43	.670

Note. B, *unstandardized regression coefficient;* β, *standardized regression coefficient.* Bs, βs, ts, *and* ps *are reported for the full model.*

TABLE 6. Regression of Peak Rating on Predictor Variables for Younger (Age <44) and Older (Age ≥44) Tournament-Rated Chess Players

	Younger (n = 226)					Older (n = 102)				
	R^2	B	β	t	p	R^2	B	β	t	p
Model	.543					.540				
Peak study time		175.7	.35	5.81	<.001		140.5	.26	2.84	.005
Peak tournament play		78.1	.13	2.32	.021		65.6	.12	1.31	.194
Peak private instruction		107.1	.13	2.66	.008		43.5	.06	0.72	.471
Peak group instruction		18.0	.03	0.47	.639		12.2	.02	0.26	.798
Books		139.9	.27	4.92	<.001		219.1	.47	5.86	<.001
Serious age		–7.7	–.14	–2.52	.013		–2.8	–.12	–1.60	.114

Note. B *= unstandardized regression coefficient;* β *= standardized regression coefficient.*

serious study time and books were significant predictors of peak rating (both positive).

DISCUSSION

The primary objective of this reanalysis was to assess how much of the variance in peak chess expertise could be explained by domain-specific experience. Domain-specific experience included a measure of the activity that may be considered deliberate practice for chess (i.e., serious study time) along with other forms of domain-specific experience. SEM revealed that, together with starting age, factors reflecting serious chess activity and chess instruction accounted for 63% of the variance in peak rating. Serious chess activity (comprising serious study time, tournament play, and books) was a stronger predictor of peak rating than chess instruction (comprising private and group instruction) (βs = .62 vs .19).

Both serious chess activity and chess instruction are presumably influenced by the resource and effort constraints described by Ericsson et al. (1993). That is, players with ample resources in the form of time, energy, and access to teachers and competitive tournaments might be able to study more, compete more often, and receive more formal instruction than those lacking such resources. Genetic factors may also contribute to individual differences in these experiential factors through gene–environment correlation. In the present context, people with high levels of various genetically influenced ability (e.g., general intelligence) and nonability (e.g., grit) traits may be more likely to seek out and persist at chess-related activities than those with lower levels of these traits. In line with this speculation, two recent studies found substantial heritability for how much individuals engaged in music practice (Hambrick & Tucker-Drob, 2015; Mosing, Madison, Pedersen, Kuja-Halkola, & Ullén, 2014). In general, both environmental and genetic factors may contribute to individual differences in serious chess activity and chess instruction.

There were four additional findings. The first was that serious starting age had a significant effect on peak rating (β = –.15), above and beyond the chess experience factors. Chess players who started playing chess seriously at a younger age tended to attain higher levels of skill than those who started later, even when amount of chess experience was controlled for. One possible explanation, summarized by Gobet and Campitelli (2007), is that there is a sensitive period for skill acquisition in complex domains such as chess, such that the development of chess expertise is facilitated by the neural plasticity that characterizes human infancy and childhood (Stiles, 2000). Although these results are consistent with the sensitive period hypothesis, further research is necessary before strong conclusions can be made about this possibility. Another explanation is that other chess experience factors mediate the relationship between serious starting age and chess skill. For example, chess-playing parents might encourage their children to start studying at a younger age than non–chess-playing parents. Children of chess players would have access to informal instruction from their parents, but this training would not be included in participants' retrospective estimates of "formal chess instruction from a teacher or trainer" (Charness et al., 2005).

One could argue that participants who started playing chess seriously at an older age may not have reached their true peak rating yet, because they did not have as much time to train as participants who started to play chess seriously at a younger age. If so, the magnitude of the relationship between serious starting age and peak rating would be inflated. Counter to this possibility, however, serious starting age significantly predicted peak rating even after domain-specific experience factors were controlled for, indicating that differing amounts of training do not account for the relationship between serious starting age and peak rating.

The second additional finding was that after serious study and formal instruction were accounted for, tournament play significantly contributed to peak rating. However, the contribution of tournament play to peak rating was very small ($\Delta R^2 = .006$) compared with the proportion of variance accounted for by serious study and formal instruction ($R^2 = .415$). This result is consistent with the claim that activities such as competition are far from optimal for learning (see Ericsson et al., 1993, but also Howard, 2013). However, some minimal amount of tournament play is clearly necessary to advance in skill, as seen in the age subsets analysis, where tournament play predicted peak rating in younger but not older players. Determining the minimal amount needed remains an open question.

The third additional finding was that there was no evidence that formal instruction increased the effectiveness of serious study on chess rating. Specifically, neither the peak serious study time × peak private instruction nor the peak serious study time × peak group instruction interaction contributed significantly to the prediction of peak chess rating. Thus, Ericsson et al.'s (2007) hypothesis that formal instruction should increase the efficacy of serious study was not supported. Our results support claims by some prominent self-taught chess players that it is possible to learn chess without the help of a coach (Charness et al., 2005), although future studies should examine whether instruction from world-class coaches increases the effectiveness of serious study. Chess books also did not moderate the effect of serious study on peak rating. However, it is possible that formal instruction or other training aids could increase the efficacy of serious study in other domains. For example, in music, a teacher may assign specific scales or études for a student to work on that target the student's weaknesses.

The fourth additional finding was that, similar to Charness et al.'s (2005) findings, the effects of predictors of peak rating differed in magnitude for older (age 44 or older) and younger (age less than 44) chess players. In particular, tournament play, private instruction, and serious age were significant predictors of peak rating for younger chess players but not for older chess players. There are a number of possible explanations for this finding. Retrospective estimates of training histories may be more accurate for younger chess players than for older chess players (see Park & Gutchess, 2005). Another possibility is that the relative importance of training activities changes as players acquire more skill. Future research should examine this possibility using longitudinal designs.

Limitations

One caveat to the present results is that they are based on results from a cross-sectional study in which participants gave retrospective estimates of practice time (and other activities). There are at least two reasons why this type of research design may lead to imprecise estimates of the correlation between practice and skill (see Hambrick, Macnamara, Campitelli, Ullén, & Mosing, 2016). The first is selective attrition: People high in an ability predictive of success in some domain may be more likely to persist and accumulate large amounts of practice than those lower in the ability (Sternberg, 1999). Findings of a longitudinal study by de Bruin, Rikers, and Schmidt (2007) have been used to argue against this possibility (see Ericsson & Towne, 2010). de Bruin et al. studied Dutch chess players selected for a national training program and found that the effect of serious study on chess skill was equal for players who persisted in the program and for those who dropped out. However, de Bruin et al.'s chess players were top performers for their age; chess players who did not reach high levels of performance were not sampled. As de Bruin, Kok, Leppink, and Camp (2014) noted, longitudinal studies that track skill development from the novice to expert level are needed to better understand the influence of selective attrition on skill acquisition research.

The second reason that practice–skill correlations may be imprecise in cross-sectional studies of expertise is potential bias in estimates of practice: To some extent, participants may base their practice estimates on their current level of skill rather than on accurate recollection of their engagement in practice. That is, unable to recall exactly how much practice they engaged in across their careers, particularly in the early stages, high-skill participants may reason that they practiced more than they did, whereas low-skill participants may assume that they practiced less than they did. Alternatively, estimates of practice could be biased in the opposite direction. For example, high-skill players who believe their success is attributable to talent may underestimate their practice time. At present, it is unclear whether the net effect of these biases inflates or suppresses the correlation between practice and skill.

Another limitation of the cross-sectional design is that we were unable to assess the relationship between changes in training and changes in rating longitudinally. We could not perform these analyses because we did not have the necessary data (i.e., changes in rating over time), and the players were deidentified, which prevented us from searching for this information using chess databases. Vaci and Bilalić (2017) recently outlined methods for longitudinal data analysis using chess databases, demonstrating how one can assess, for example, age-related declines in skill, birth cohort effects, and differences between male and female chess players using openly available data. This type of study is an important goal for future research (for examples, see the recent special issue of *Topics in Cognitive Science* [Gray, 2017] and also Stafford & Dewar, 2014).

A final limitation of this study is that we could consider only variables that were measured in the original Charness et al. (2005) project. For example, it is possible that dispositional variables such as effortful control or self-regulation could explain some of the shared variance between training time and chess skill, because both serious study and tournament competition require prolonged periods of concentration. If so, the estimated effect of serious study on chess rating would be upwardly biased in the present analyses by the omission of these predictor variables. Similar arguments could be made for cognitive ability, if people higher in cognitive ability are more likely to study chess (see Sala et al., 2017). A comprehensive analysis of chess skill incorporating all potentially relevant predictor variables is a worthwhile goal for future research.

Conclusions

The results of this study affirm that deliberate practice and other forms of domain-specific experience are important pieces of the expertise puzzle. Indeed, all measures of domain-specific experience combined accounted for 63.4% of the variance in peak rating. Serious study, which has been argued to meet the description of deliberate practice, was a strong predictor of peak rating in multiple regression analyses (βs ranging from .26 to .45). Contrary to arguments made by proponents of the deliberate practice framework (e.g., Ericsson et al., 1993), however, individual differences in deliberate practice left a substantial proportion of the variance in peak chess rating unexplained. That is, in multiple regression analyses, serious study accounted for less than half of the variance in peak rating. This finding is consistent with results of Hambrick et al.'s (2014) meta-analysis, which showed that although deliberate practice is an important predictor of chess skill, it left 66% of the variance in expertise unexplained after measurement error was corrected for.

The results of the structural equation model analysis suggest that other unmeasured factors must matter, too. To be exact, all measured domain-specific experience left 36.6% of the variance in peak rating unexplained. This result is consistent with Charness et al.'s (2005) finding that approximately 60% of the variance in peak rating remained unaccounted for by training. The fact that we accounted for approximately 20% more of the variance in peak rating than Charness et al. (2005) is probably because structural equation modeling corrects for unreliability, and could be in part due to the variables included in the models. For example, we included serious starting age and the number of chess books owned, whereas Charness et al. (2005) did not.

What accounts for this remaining variance is a critical question for future research. To begin to answer this question, future studies should include measures of cognitive ability, informal instruction, coaching quality, and the use of other training aids such as online databases. Longitudinal analyses could also be used, for example, to determine whether the effect of formal instruction on serious study differs depending on the age at which formal instruction begins. An equally important challenge is to characterize and assess the microstructure of training and practice activities that help build the cognitive mechanisms (e.g., knowledge base of chess patterns, search processes) supporting expert performance.

NOTES

Address correspondence about this article to Alexander P. Burgoyne, Michigan State University, Psychology Building, 316 Physics Road, Suite 48, East Lansing, Michigan 48824 (e-mail: burgoyn4@msu.edu).

1. This rating gives points to and ranks chess players based on their tournament games and has been used by the International Chess Federation since 1971. Players with more than 2000 points are typically considered chess "experts," whereas players with less than 800 points are considered "beginners."

2. We conducted another set of analyses for participants who had no missing values. The overall pattern of results was almost identical to the results reported in the present study. All conclusions are the same. See Supplemental Materials.

3. Tournament play was not significant when current rating was examined; see Supplemental Materials at https://www.press.uillinois.edu/journals/ajp/media/chess_skill

REFERENCES

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.

Burgoyne, A. P., Sala, G., Gobet, F., Macnamara, B. N., Campitelli, G., & Hambrick, D. Z. (2016). The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. *Intelligence, 59,* 72–83.

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.

de Bruin, A. B., Kok, E. M., Leppink, J., & Camp, G. (2014). It might happen in the very beginning. Reply to Ericsson. *Intelligence, 45,* 107–108.

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

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological),* 1–38.

Elo, A. E. (1978). *The rating of chess players, past and present*. New York, NY: Arco Publishing.

Ericsson, K. A. (2005). Recent advances in expertise research: A commentary on the contributions to the special issue. *Applied Cognitive Psychology, 19,* 233–241.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49,* 725.

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

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

Ericsson, K. A., & Towne, T. J. (2010). Expertise. *Wiley Interdisciplinary Reviews: Cognitive Science, 1,* 404–416.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4,* 272.

Flather, M. D., Farkouh, M. E., Pogue, J. M., & Yusuf, S. (1997). Strengths and limitations of meta-analysis: Larger studies may be more reliable. *Controlled Clinical Trials, 18,* 568–579.

Gobet, F., & Campitelli, G. (2007). The role of domain-specific practice, handedness, and starting age in chess. *Developmental Psychology, 43,* 159.

Gray, W. D. (2017). Game-XP: Action games as experimental paradigms for cognitive science. *Topics in Cognitive Science, 9,* 289–307.

Hambrick, D. Z., Macnamara, B. N., Campitelli, G., Ullén, F., & Mosing, M. A. (2016). Chapter one: Beyond born versus made: A new look at expertise. *Psychology of Learning and Motivation, 64,* 1–55.

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence, 45,* 34–45.

Hambrick, D. Z., & Tucker-Drob, E. M. (2015). The genetics of music accomplishment: Evidence for gene–environment correlation and interaction. *Psychonomic Bulletin & Review, 22,* 112–120.

Hensch, T. K. (2004). Critical period regulation. *Annual Review of Neuroscience, 27,* 549–579.

Howard, R. W. (2013). Practice other than playing games apparently has only a modest role in the development of chess expertise. *British Journal of Psychology, 104,* 39–56.
Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60–99.
Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford.
Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice, 2,* 59–67.
Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions a meta-analysis. *Psychological Science, 25,* 1608–1618.
Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R., & Ullén, F. (2014). Practice does not make perfect no causal effect of music practice on music ability. *Psychological Science, 25,* 1795–1803.
Park, D. C., & Gutchess, A. H. (2005). Long-term memory and aging: A cognitive neuroscience perspective. In R. Cabeza, L. Nyberg, & D. C. Park (Eds.), *Cognitive neuroscience of aging: Linking cognitive and cerebral aging* (pp. 218–245). Oxford, UK: Oxford University Press.
Roring, R. W., & Charness, N. (2007). A multilevel model analysis of expertise in chess across the life span. *Psychology and Aging, 22,* 291.
Sala, G., Burgoyne, A. P., Macnamara, B. N., Hambrick, D. Z., Campitelli, G., & Gobet, F. (2017). Checking the "Academic Selection" argument. Chess players outperform non-chess players in cognitive skills related to intelligence: A meta-analysis. *Intelligence, 61,* 130–139.
Stafford, T., & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological Science, 25,* 511–518.
Sternberg, R. J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology, 24,* 359–375.
Stiles, J. (2000). Neural plasticity and cognitive development. *Developmental Neuropsychology, 18,* 237–272.
Vaci, N., & Bilalić, M. (2017). Chess databases as a research vehicle in psychology: Modeling large data. *Behavior Research Methods, 49,* 1227–1240.