Intelligence xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Intelligence

Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms ☆

K. Anders Ericsson

Department of Psychology, Florida State University, Tallahassee, FL 32306-1270, USA

ARTICLE INFO

Article history:
Received 1 December 2013
Accepted 1 December 2013
Available online xxxx

Keywords:
Expert performance
Deliberate practice
Long-term working memory
Innate talent
IQ

ABSTRACT

Many misunderstandings about the expert-performance approach can be attributed to its unique methodology and theoretical concepts. This approach was established with case studies of the acquisition of expert memory with detailed experimental analysis of the mediating mechanisms. In contrast the traditional individual difference approach starts with the assumption of underlying general latent factors of cognitive ability and personality that correlate with performance across levels of acquired skill. My review rejects the assumption that data on large samples of beginners can be extrapolated to samples of elite and expert performers. Once we can agree on the criteria for reproducible objective expert performance and acceptable methodologies for collecting valid data. I believe that scientists will recognize the need for expert-performance approach to the study of expert performance, especially at the very highest levels of achievement.

1. Introduction

Contributors to the special issue on “Development of Expertise” criticized Ericsson, Krampe, and Tesch-Römer's (1993) theoretical framework for studying expert performance. In this response I will propose that many of these unflattering descriptions are due to the fact that the Expert-Performance framework is qualitatively different from the traditional theoretical frameworks for relating individual differences in ability. The fundamental controversy concerns the relation between normal and beginning levels of performance and the attainable expert levels of performance. The researchers working within the individual differences framework argue that the structure of expert performance can be extrapolated from the performance–ability relations observed in the general adult population. The structure of expert performance merely corresponds to extreme cases on the underlying ability distributions, and thus ability differences in general cognitive ability remain predictive of performance at the highest level. In contrast, the expert-performance framework hypothesizes that new cognitive mechanisms are gradually acquired during the extended period and they mediate the superior performance, thus leading to qualitative differences in structure compared to untrained performance.

1.1. The study of the acquisition of expert performance through deliberate practice

During the last 20 years since our article in 1993 (Ericsson et al., 1993) peoples' conceptions about expertise and expert performance have changed. In the early 1990s, when our paper was written, the contemporary theories of skill acquisition (Anderson, 1982; Fitts & Posner, 1967) at that time proposed accounts for the acquisition of everyday skills such as driving a car, typing, and navigating in an unfamiliar environment. For these skills most individuals had the primary goal of

☆ Author notes: I am grateful for the financial support provided by the FSCW/Conradi Endowment Fund of Florida State University Foundation. The author wants to thank Walter Boot, Edward Cokely, Len Hill, Colleen Kelley, and Jerad Moxley for their valuable comments and suggestions on earlier drafts of this article.

E-mail address: ericsson@psy.fsu.edu.

0160-2896/$ – see front matter © 2013 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.intell.2013.12.001

reaching a level of proficiency that allowed them to perform these tasks adequately and effortlessly. At that time typical laboratory studies of skill acquisition were limited to 1 or 2 h per study and hardly ever lasted for more than 50 h total. A prototypical study involved choosing one of 4–16 alternatives. The task needed to be simple enough to be explained to college students in less than 15 min by reading instructions followed by a short period of warm-up. In their review of skill acquisition Fitts and Posner (1967) summarized the general structure of this type of skill acquisition and Ackerman (1987) proposed how individual differences in basic abilities would be predicted to correlate with performance at the different stages. During the first introductory phase of learning the skill (Fitts & Posner, 1967), beginners try to understand the requirements of the activity and focus on generating adequate actions while avoiding gross mistakes. During this phase individual differences in participants' performance were correlated with their general cognitive abilities (Ackerman, 1987). In the second phase of skill acquisition, once people had accumulated more task experience, salient mistakes become increasingly rare, sequences of actions are generated in a smoother fashion, and learners no longer need to concentrate to maintain an acceptable performance. Individual differences in participants' performance during this phase correlate with their tested spatial ability (Ackerman, 1987). During the third phase of learning individuals' performance skills become increasingly automated, and they are able to execute their skills smoothly and with minimal effort. As a consequence of automatization, performers lose their ability to control the execution of those skills, which makes intentional modifications and adjustments difficult. In the automated phase of learning, performance reaches a stable plateau, and no further improvements are typically seen. When the performance is triggered automatically, thus short-circuiting the cognitive and spatial representations, the individual differences in performance will primarily reflect the motor processes and the associated motor abilities (Ackerman, 1987). Drawing on Shiffrin and Schneider's (1977) theory for automation of consistent reactions Ackerman (1987) proposed that during acquisition of these types of skills the cognitive and spatial factors were eventually short-circuited and individual differences in these basic abilities would not correlate with resulting performance measured by the speed of the responses.

In contrast to the acquisition of everyday skills Ericsson et al. (1993, p. 363) proposed the study of objective reproducible performance of "exceptional individuals, whose performance in sports, the arts, and science is vastly superior to that of the rest of the population". We found that acquiring these high levels of performance required years and even decades of demanding practice, a finding that has now been replicated in every type of expert performance that has been studied to date. We argued that with the extended periods of intense practice there is room for circumvention, adaptation and/or fundamental change in basic abilities: "the perceptual and motor systems show great adaptability in response to extended practice (a phenomenon discussed earlier in this article), it may be inappropriate to generalize the findings from relatively simple tasks involving 2–20 h of practice to expert performance acquired during a 10-year period of intense preparation" (p. 396). Expert performers in music, chess, and sports are constantly learning new things. Thus they maintain cognitive control over their performance and their performance does not tend to become fully automated. Ericsson et al. (1993) commented on how expert performers acquire domain-specific memory skills to recall past game situations, plan future actions, and evaluate their current performance. As is illustrated in Fig. 1, Ericsson et al. (1993) argued that expert performance requires the acquisition of new cognitive structures to enhance domain-specific performance and that "experts can acquire cognitive skills enabling them to circumvent the limits of short-term memory capacity and serial reaction time. This research rules out the hypothesis that individual differences in those functions will influence and constrain final adult performance" (p. 396). In contrast to many proposals that individual differences in working memory are limited by a basic general capacity for transient short term storage (Baddeley, 1986), Ericsson and Kintsch (1995) showed that expert performers develop long-term working memory (LTWM), where information is rapidly stored in long-term memory (LTM) associated with retrieval cues that allow the expert to access this information efficiently whenever the information is relevant for processing. Given that these skills are constructed based on available knowledge in LTM and choices about encoding methods, there will be substantial individual differences in the structure of the acquired skills, which have been demonstrated with experimental methods especially for memory experts (Ericsson & Kintsch, 1995; Hu & Ericsson, 2012). Investing a large proportion of one's lifetime in challenging deliberate practice activities causes profound qualitative changes in physiology and psychology.

1.2. The definition of expert performance

Because expert performance is qualitatively different than other types of human performance, it requires the study of reproducibly superior performance on representative tasks that capture the essence of expertise in real world domains.

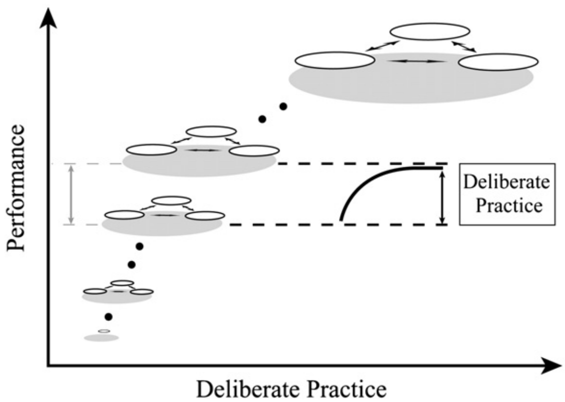

Fig. 1. A schematic illustration of the acquisition of expert performance as a series of states with mechanisms of increasing complexity for monitoring and guiding future improvements of specific aspects of performance.
Adapted from "The scientific study of expert levels of performance can guide training for producing superior achievement in creative domains" by K.A. Ericsson in Proceedings from *International conference on the cultivation and education of creativity and innovation* (p. 14). Beijing, China: Chinese Academy of Sciences. Copyright 2009 by International Research Association for Talent Development and Excellence.

Most commentators seem to accept our definition of expert performance and at least some of its implications. For example, Plomin, Shakeshaft, McMillan, and Trzaskowski (2014-this issue) acknowledge that: “It is important to understand the origins of expertise as it exists in the real world of sports, arts and skills” (p. 3). Ironically, however, Plomin et al. (2014-this issue) then go on to study precocious performance of 12-year-old children, who perform at a very high level on a reading test compared to their age-matched peers, that nearly everyone eventually masters within a few years. However, this level of performance is only precocious and six years later the majority of all of the other students are able to match their performance. In this response I will use the definition of Ericsson and Lehmann (1996, p. 277) and refer to expert performance “as consistently superior performance on a specified set of representative tasks for a domain” without any age conditions. Unfortunately other investigators use the term differently. For example, Wong, Palmeri, and Gauthier (2009) acknowledge that real-world expertise may differ but that they can produce performance showing “perceptual expertise” for classifying artificial figures after 2–10 h of training and call these participants “Ziggerin experts”. Other investigators refer to children with more knowledge about dinosaurs as experts (Gobbo & Chi, 1986). Finally, the requirement that experts display measureable superior performance in their domain is not met by many types of expertise. For example, Wai (2014-this issue) does not specify the nature of the reproducibly superior objective performance that his senators and billionaires are able to demonstrate in comparison to their peers. In many domains individuals referred to as experts by their peers are often unable to perform at a superior level to their peers on representative tasks from the domain (Ericsson, 2006a, 2009; Ericsson & Lehmann, 1996).

All papers in this issue criticize my earlier research for attributing too much emphasis to the effects due to deliberate practice. In our original paper, Ericsson et al. (1993) were very explicit that there might be other types of individual differences than those linked to innate talent: “It is quite plausible, however, that heritable individual differences might influence processes related to motivation and the original enjoyment of the activities in the domain and, even more important, affect the inevitable differences in the capacity to engage in hard work (deliberate practice)” (p. 399). The ability to engage in deliberate practice is an obvious requirement for improving performance through deliberate practice, but not all individuals may be able or willing to do so.

I find it disappointing that so many contributors seem to have misinterpreted quotes, where we argued that deliberate practice was necessary and based on our original study could explain “the *major* facts about the nature and scarcity of exceptional performance” (Ericsson et al., 1993, p. 392, italics added). They also seem to have overlooked many paragraphs where we tried to make our views on innate talent as clear as possible, such as this quote from our target paper in the *High Ability Studies*' issue on “Giftedness and Expertise”—a paper that most extensively addressed the issues of the special issue of *Intelligence*, but that Hambrick et al. (2014-this issue) seem unaware of and therefore did not cite:

“A common misconception of the expert performance framework is that this approach denies the possibility that differences in innate talent could ever be able to explain individual differences in attainable performance. The expert performance framework merely requires that valid evidence for innate talents must be presented and reviewed before it is accepted. This framework has long acknowledged the possibility that individual genetic differences might causally explain individual differences in elite achievement. However, according to recent reviews (Ericsson, 2007a, 2007b, 2007c) no evidence currently exists, with the exception of height and body size” (Ericsson, Roring, & Nandagopal, 2007, pp. 40–41).

In my response I will pursue this same argument as I examine the proposed empirical evidence in support of innate abilities and talent that would constrain, and thus correlate with the performance that an individual can attain in a domain of expertise. The central goal of Ericsson et al. (1993) and this response remains to identify immutable constraints on the acquisition of various types of expert performance that cannot be overcome or circumvented by the most effective forms of practice (deliberate practice) in order to provide “unique evidence on the potential and limits of extreme environmental adaptation and learning” (p. 363).

1.3. Outline of my response

In the second section of the main body of my reply I will review the cited studies reporting significant relations between performance on tests of cognitive ability and domain-specific performance (Ackerman, 2014; de Bruin, Kok, Leppink, & Camp, 2014-this issue; Grabner, 2014-this issue; Hambrick et al., 2014-this issue). I will show that the findings are consistent with our proposal for the acquisition of expert performance, where acquired mechanisms gradually circumvent the role of any basic general cognitive capacities and thus reduce and even eliminate significant relations between general cognitive ability and domain-specific performance at the expert level of performance.

In the third section I will argue that there are no available estimates on heritability for attained expert performance in twins, when we define expert performance as reproducibly superior performance in a domain of expertise independent of age. I will also show that the other evidence cited by Plomin et al. (2014-this issue) for genes being necessary for attaining expert performance is currently lacking, with the long established exception of genetic effects on height and body size.

In the fourth section I will review the claim that precocious performance on Scholastic Achievement tests in Mathematics (SAT-M) in the 8th grade is predictive of expert adult performance in mathematics and science (described by Wai, 2014-this issue). I will show that these findings can be accounted for within the expert-performance framework without acknowledging that exceptional innate mental capacities are necessary to attain expert levels of performance.

In the fifth section I will discuss the relation between self-reported amounts of practice and current domain-specific performance. The expert-performance approach is consistent with Hambrick et al.'s (2014-this issue) first conclusion namely that the reported amount of practice is significantly correlated with current level of performance in music and chess. However, Hambrick et al.'s (2014-this issue) second conclusion concerns the relation of attained performance and the total

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

amount of practice during the individuals' past life. Their analysis ignores the effects of forgetting, injuries, and accidents, along with the differential effects of different types of practice at different ages and levels of expert performance.

In a sixth and final section I will more briefly discuss a few remaining issues.

2. Issue 1: General cognitive ability and the attained level of performance

There are two types of processes that might give rise to correlations between performance on tests of cognitive ability and performance in a given domain, which Ackerman (2013) in his recent chapter on "Expertise and Intelligence" mentions explicitly. The first process concerns selection processes, where there are restricted opportunities and limited resources to engage in the activity and associated training. Nobody can deny that the necessary training to succeed in domains of expertise, such as sailing, polo, and Formula One car racing, would require financial resources that are beyond the scope of most children and adolescents and their respective families. Access to books, programmable computers, and videos of expert performers may facilitate early development of some types of expert performance. In his famous studies of international level performers Bloom (1985) found that the family had invested a lot of time, money, and emotional support, so we would expect a correlation with Socio-Economic Status (SES). Many undergraduate and graduate training programs are quite selective in the acceptance of students into their programs and require superior high-school grades and test scores, such as the Scholastic Aptitude Test (SAT). Performance on SAT is estimated to correlate with measures of general intelligence (g) (beta = 0.78) (Coyle, Purcell, Snyder, & Kochunov, 2013).

There might be a second type of relation between cognitive ability and performance. It may be that cognitive ability is correlated with performance uniformly across the range of skill starting with beginners all the way to professionals. The expert-performance approach proposes that performance on tests of general cognitive ability will be correlated for beginners—a finding consistent with laboratory studies (Ackerman, 1987) and reviews of job performance (Schmidt & Hunter, 2004). For individuals who have acquired cognitive structures that support a high level of performance the expert-performance framework predicts that these acquired cognitive structures will directly mediate superior performance and thus diminishing correlations between general cognitive ability and domain-specific performance.

2.1. Group differences in cognitive ability between participants and controls

Consistent with admissions criteria for entering research universities as an undergraduate and, in particular, as graduate students, scientists score much higher on tests of intelligence than the general population. A number of early studies showed that eminent scientists had higher than average IQ's (Roe, 1953a, 1953b) ranging from 121 to 177, a larger group of university scientists ranged between 110 and 141 (N = 131, M = 126.5, SD = 6.2) (Gibson & Light, 1967); and over 5000 science doctorial graduates recorded a mean of 130.8 (SD = 17.1 on the Army Standard Scale, which has M = 100, SD = 20) (Harmon, 1961). Law, Wong, Huang, and Li (2008) found that their sample of scientists had significantly higher performance on the Wonderlic Personnel Test (N = 102, M = 37.5, SD = 5.0) than undergraduate college students in a sample collected by McKelvie (1989) (N = 290, M = 26.7, SD = 6.2).

2.1.1. Football

When colleges recruit football players to their teams, these players are often selected differently from the regular students. Most universities are willing to admit strong football players even if they have a lower score on the admissions tests, such as SAT. Although most football players have attended college there are also paths to professional football playing that does not involve playing for a college team. Lyons, Hoffman, and Michel (2009) analyzed the Wonderlic IQ score of those athletes participating in the National Football League (NFL) draft and who were selected to play in the NFL (N = 762). The average of this sample of athletes (M = 21.04, SD = 7.15) did not differ from the general population norms (M = 21.75, SD = 7.6).

2.1.2. Music

The selection of musicians differs for different types of musicians. Ruthsatz, Detterman, Griscom, and Cirullo (2008) tested high school students playing for the school band and found that their mean score on the Advanced Raven's Progressive Matrices was (M = 21.7, SD = 5.4), which did not differ from the general population with the same age (M = 19, SD = 6). This average score on the Raven was comparable to an IQ of 105. Ruthsatz et al. (2008) also tested more skilled music students, who'd been accepted to conservatories and university-based music schools and thus subjected to academic selection. These advanced musicians scored significantly higher (M = 25.2, SD = 5.7) on the Raven than the band members. The Raven scores of the skilled music students corresponded to an average IQ equivalent of 113. In support of the role of academic selection Helmbold, Rammsayer, and Altenmuller (2005) found that a group of music majors did not differ in intelligence from other university students.

2.1.3. Chess

Several contributors to the special issue (de Bruin et al., 2014-this issue; Grabner, 2014-this issue; Hambrick et al., 2014-this issue) reported that samples of active young chess players tend to have higher than average IQs. Both Frydman and Lynn's (1992) chess players and Bilalić, McLeod, and Gobet's (2007) players were around 11 years old and had average IQs of 121 and 122 respectively. Samples of adult players had average IQs of 118 (Grabner, Neubauer, & Stern, 2006), 114 (Grabner, Stern, & Neubauer, 2007), and 107 (SD = 7.5) (Doll & Mayr, 1987) When, however, Untertrainer, Kaller, Hasband, and Rahm (2006) identified a group of chess players and then found a matching control group by matching for age and education they did not find a significant difference on Raven's matrices.

In studies of other types of abilities Waters, Gobet, and Leyden (2002) found that a sample of adult chess players performed significantly worse than 550 naval recruits on a test of visuo-spatial ability. Unterrainer, Kaller, Halsband, and Rahm (2006) found that their sample of chess players performed significantly better than their control group on a

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

task assumed to measure planning ability, namely Tower of London. However, in a more recent study Unterrainer, Kaller, Leonhart, and Rahm (2011) were unable to replicate the superiority of the chess players in two experiments, providing a compelling alternative account that related to the higher motivation of chess players to perform in a competitive situation.

2.1.4. GO

The Japanese game GO is particularly interesting in that no computer program has yet even come close to competing with human masters. Lee et al. (2010) collected data on IQ from 16 Baduk (Korean name for GO) experts and 19 age- and sex-matched controls, who did not play Baduk or chess. The Baduk experts had an average IQ ($M = 93.2$, $SD = 10.4$), which was almost significantly ($p = 0.052$) *lower* than the average of the control group ($M = 101.2$, $SD = 13.1$). In a more recent study Jung et al. (2013) found that a group of 17 Baduk experts had an average IQ of 93.1 ($SD = 10.1$). The pattern of means for expert GO players compared to control populations is the reverse of that observed for chess players in USA and Europe compared to control populations.

In sum, the general pattern of group means is consistent with selection effects rather than consistent benefits of higher intelligence for individuals engaged in the corresponding domains.

2.2. Relation between general cognitive ability and performance among participants

Based on the expert performance framework the correlation between domain-specific performance and tests of cognitive ability is hypothesized to differ as a function of the level of acquired skill (see Fig. 1). It is hard to assess this hypothesized pattern because most contributors to this special issue do not differentiate studies of beginners, expert performers, as well as mixed samples of beginners, intermediates and experts. To evaluate this hypothesized pattern, I will discuss the findings in the three types of studies in separate sub-sections.

2.3. Studies of beginners and less-accomplished performers

According to the expert-performance framework, beginners would have to rely on their pre-existing traits and abilities and thus would be predicted to show correlations between domain-related performance and performance on tests of their pre-existing abilities, such as general cognitive abilities. In many domains beginners can benefit from being able to read and comprehend instructions and books, which corresponds to an ability assumed to correlate with scores on tests of verbal abilities and intelligence.

2.3.1. Chess

The new study reported by de Bruin et al. (2014-this issue) provides a very nice example of how IQ is correlated with chess performance in a sample of beginning young chess players. The two cited studies of beginning chess players, namely children playing competitive chess, uncovered different patterns of results. When Frydman and Lynn (1992) split their sample into the third with the highest chess rating, the average third and the third with the lowest chess ratings, they found no difference in their average IQ scores with 122, 123, and 117, respectively. In contrast, Bilalić et al. (2007) found that their most highly rated children had a significantly higher IQ ($M = 133$) than the rest of their sample ($M = 114$). Based on adult standards for criteria for chess experts these children were not playing chess at that level. In fact, the average chess ratings for these children were typically below 1500.

2.3.2. Music

When students begin to play a musical instrument their ability to perform pieces of music is very limited. Ruthsatz, Ruthsatz, and Stephens (2014-this issue) cites Shuter's (1968) review of 65 music studies to support the claim of "a positive correlation between musical achievement and general intelligence of .35" (p. 2). These studies predominantly study children in the age range of 7–12, who have had music training or music training for less than a year or so and test memory and perception of music rather than music performance. For example, Ruthsatz et al. (2014-this issue) mention explicitly a study by Lynn and Gault (1986) showing "a positive correlation between musical achievement and general intelligence" (p. 2). A more careful reading of Lynn and Gault's (1986) paper shows that they studied all of the 93 children in an elementary school between ages 9–11 and did not test their ability to perform, but rather gave them tests of memory for music, pitch discrimination and detecting the number of notes played in a chord on the piano. In a follow-up paper Lynn, Wilson, and Gault (1989) reported analyses of the same sample and an analysis of 210 10-year old children given the same tests of perception and memory. The research cited by Ruthsatz et al. (2014-this issue) shows essentially that there is a correlation between performance on tests of general intelligence and the performance on memory and perceptual tasks involving music stimuli in the 0.3 to 0.4 range for children without a history of music training. This line of research was motivated by a strong belief in the existence of measureable innate music talent that could be used to guide young "talented" children to early music instruction.

Even after only six months of instruction it is possible to measure the attained music performance. For example, Young (1971) tested 91 5th grade students who had received piano instruction for around six months in a group setting with additional opportunities to train. During their final examination they had to play one piece practiced during the two weeks prior to the test along with sight-reading and improvisation. Young (1971) found a correlation between the rated performance and intelligence $r(89) = .38$, $p < .001$. Other studies have shown correlations between judged performance of band members (4th to 6th grade middle-school students) and intelligence (Hufstader, 1974) and between criterion-based performance of 5th grade students in their first year of music instruction and achievement in reading in regular school (Klinedinst, 1991).

Studies have examined correlations between intelligence and objective music performance as a function of the number of years that students in a music school had played their instruments. For example, Young (1971) reports that "an intelligence test score has predictive validity that is inversely

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

proportional to number of years of study" (p. 397) from $r(58) = .49$, $p < .01$ in the first year to $r(48) = .24$, n.s. in the 6th year. Finally, several investigators have found correlations between music performance and IQ-related measures among high school band members. Ruthsatz et al.'s (2008) studied high school band members and found a significant correlation between IQ and music performance ($r(176) = .25$, $p < .01$). Gromko (2004) found a significant correlation between sight reading, on the one hand, and reading comprehension and mathematics performance.

Consistent with the predictions from the expert-performance framework, there are consistently significant correlations between performance on tests of general cognitive abilities and domain-specific performance for beginners and less accomplished individuals, typically children and adolescents.

2.4. Studies of experts and highly-accomplished performers

The crucial difference between the expert-performance framework and the individual difference framework concerns predictions about correlations between tests of cognitive ability administered before start of training and attained domain-specific performance. If acquired cognitive skills have circumvented any constraining general cognitive capacities then correlations would be low and non-significant and if these correlations are preserved among expert performers then these abilities continue exerting their effects on performance.

2.4.1. Scientific research

Fifty years ago MacKinnon (1962) studied architects differing in creativity. He used a procedure involving editors of architecture journals and ratings by expert architects to identify a sample of highly creative architects. He found a low and insignificant correlation ($r(38) = -.08$, $p > .05$) between performance on an intelligence test and creativity, where the highly creative group's IQ ranged widely on the Terman Concept Mastery Test. He concluded that "It is clear, however, that above a certain required minimum level of intelligence which varies from field to field and in some instances may be surprisingly low, being more intelligent does not guarantee a corresponding increase in creativeness" (p. 288). Studies of the relation between creative and less creative scientists and their cognitive abilities failed to reveal significant differences in smaller groups (Walker, 1955). Harmon (1961) collected a large data base of IQ scores of students attaining science doctorates by contacting their high school. Bayer and Folger (1966) extracted IQ scores from a set of 228 biochemists and related these scores to the number of citations (an index of professional success) and found an insignificant correlation ($r(226) = -.049$, $p > .05$). The corresponding correlations between IQ and number of publications with any citations ($r(226) = -.027$, $p > .05$) and average number of citation per publication ($r(226) = -.023$, $p > .05$) were also not significant. Most interestingly they found highly significant correlations between the prestige of the university awarding the doctoral degree for number of citations ($r(226) = .214$, $p < .001$), number of articles cited ($r(226) = .182$, $p < .001$) and average number of citation per article ($r(226) = .226$, $p < .001$). In a replication of Bayer and Folger (1966) Cole and Cole (1973) studied "499 academic physical, biological, and social scientists" (p. 69) and they found that number of published papers was not correlated with IQ ($r(497) = .05$, $p > .05$) nor was the number of citations significantly correlated with IQ ($r(497) = .06$, $p > .05$). Interestingly they did find a correlation between IQ and the prestige of the department, where the scientist was working at the time ($r(497) = .27$, $p < .001$). After conducting a number of analyses, Cole and Cole (1973) concluded: "In short, these data offer preliminary evidence that although innate ability (as measured by I.Q.) is not correlated with the quality of research role-performance, it is in some way recognized."(p. 70). They suggest that "Verbal scientists with high I.Q.'s may be able to convince their colleagues that their work is of high quality than a completely objective evaluation might indicate" (p. 70). These studies are particularly valuable as they test IQ in high school, rather than testing middle–aged scientists after they've made significant contributions to their field.

Similarly, Roe (1953a, 1953b) found that many eminent scientists had IQ's below the average for PhD's and concluded it is "not essential to have this ability *at the highest level* in order to become an eminent scientist" (p. 164). She found that eminent scientists in different academic disciplines had different pattern of strengths on the different subtests concluding "how well you do in the field is partly a function of your capacity for that particular field, but even more a function of how hard you work at it" (p. 170). Roe (1953a, 1953b) tested her eminent scientists between the ages of 38 and 60, so it is quite possible that the tests reflected, at least to some degree, knowledge and abilities acquired during the 20-year period as active scientists, so the pattern of different abilities might, at least to some degree, be a consequence of focused engagement in a field of study for decades rather than an innate pre-requisite for future success. Some Nobel Prize-winning scientists have reported their IQ scores and Root-Bernstein et al. (2008) collected a list verified by biographies: Richard Feynman's IQ was 126, James Watson's IQ was 124, and William Shockley's IQ was 125. The longitudinal study of 40 scientists conducted by Eiduson (1962) contained four Nobel Prize winners and 11 member of the National Academy of Science along with much less successful scientists. When Root-Bernstein, Bernstein, and Garnier (1993) analyzed their scores on Miller's analogies, there was no significant correlation between scientific productivity and impact and scores on that test of reasoning. More recently, Law et al. (2008) found a non-significant correlation ($r(100) = -.07$, $p > .05$) between measured intelligence (Wonderlic Personnel Test) and a formal evaluation based on the scientists' "current and past research outputs" (p. 61).

2.4.2. Chess

The only study of chess masters with a rating of 2200 or better was conducted by Doll and Mayr (1987). They did not find a correlation between IQ and chess rating for the 27 players with an average chess rating ($M = 2301$, $SD = 54$) compared to the average chess rating ($M = 1500$, $SD = 200$). Given that the IQ for this group was ($M = 106.5$, $SD = 7.5$) both variables have less variability than the average population, and thus range restriction might be a contributing factor. They then calculated the changes in chess rating over one year for a subset of players, namely 1984/1985 ($N = 16$, $M = 5.6$, $SD = 27.6$) and 1985/1986 ($N = 17$, $M = 3.2$, $SD = 15.3$)

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

and calculated correlations for the overall IQ for changes in 1984/1985 ($r(14) = .04$, $p > .05$) and in 1985/1986 ($r(15) = .41$, $p > .05$). They also calculated 16 correlations with all the 8 sub-tests and correlations ranged between ($r(14) = -.07$) to $r(15) = .59$, but using a Bonferroni correction, none of correlations with the subtests is significant.

How important is intelligence for expert and elite chess performance? Perhaps the most informative case study is the intelligence testing of former World Champion Gary Kasparov by reporters at Der Spiegel (1987). His IQ was estimated at 120 based on the Raven test, which is very close to the average of all chess players (Howard, 2008)—thus not very predictive of world-class chess performance. It is also interesting to note that some chess masters had surprisingly low IQs. Grabner et al. (2007) report that one chess master with a rating close to 2400 had an IQ of around 80 and the 99% confidence intervals for individual chess masters' IQs in Doll and Mayr's (1987) study ranged from 86 to 127. Furthermore, Grabner et al. (2007) explicitly rejects the idea that exceptionally high scores on IQ-tests are associated with very high levels of chess performance: "When it [high chess playing ability] is defined as an ELO ranking above 2200 (advanced players or experts, cf. Charness, Tuffiash, Krampe, Reingold, & Vasyukova 2005; i.e., 7% of the sample), for verbal and numerical IQs the threshold seems to lie somewhat higher (at about 110–115). Interestingly, the scatterplots also show that the highest-rated participants in the present sample are not those with the highest verbal or numerical IQs" (Grabner et al., 2007, p. 408).

2.4.3. Music

Ruthsatz et al. (2008) studied advanced music students and found non-significant relations between IQ and domain-specific performance: the university majors with intermediate levels of practice ($r(17) = .24$, $p > .05$) nor for the conservatory musicians with most practice ($r(62) = .12$, $p > .05$). The variation in cognitive ability for the expert sample ($SD = 5.7$) is not significantly smaller than it was for less skilled sample of high school band members ($SD = 5.4$).

2.4.4. Football

Hambrick et al. (2014-this issue) reports that for football players, in particular quarterbacks, selected to play in NFL there is no significant correlation between their subsequent performance in NFL and their Wonderlic IQ scores collected at the NFL draft (Lyons et al., 2009). In fact, there were no significant correlations for any of the positions on a football team with an overall correlation between football performance and IQ ($r(760) = -.04$, $p > .05$). Lyons et al. (2009) also corrected for playing time and found that the partial correlation was ($pr = .01$). In a subsequent study focused on 121 quarterbacks selected to play in the NFL Berri and Simmons (2011) analyzed both the factors associated with order of selection within the draft as well as the crucial subsequent performance in NFL. They found that Wonderlic score was significantly predictive of being selected earlier in the draft and write: "we find that the taller, smarter, faster quarterbacks who play at Division I-A schools are likely to be picked higher in the draft" (Berri & Simmons, 2011, p. 45). In an analysis of factors predicting actual performance in the NFL they found that the Wonderlic score was significant for the first year of playing, but the correlation was negative—"In other words, higher Wonderlic scores were associated with lower levels of performance" (Berri & Simmons, 2011, p. 47). Taken together these findings suggest that coaches tend to pay attention to the Wonderlic scores when they select quarterbacks, although this measure is either not significantly or significantly negatively related to subsequent performance in the NFL.

The majority of studies in this subsection with reportable correlations have a negative sign and the weighted mean correlation across the studies is $r = -.006$ based data from 1714 participants. The standard deviations associated with the IQ measures for these samples are comparable to other control populations and the range for the dependent variable of performance is appropriately large for assessing valid correlations (see Ackerman's (2014) concerns for restriction of range). Let it be clear that I am not claiming that correlation between domain-specific performance and general cognitive ability is exactly zero!! My current conclusion is that these studies have *not yet* established the fact that the attainable level of domain-specific performance is predictable from scores from tests of general cognitive ability.

2.5. Studies of performers ranging widely in skill

In two preceding sections we found that performance on tests of cognitive ability was positively correlated with performance on domain-related tasks for beginners and less-skilled performers, but found no consistent relation for highly skilled individuals. When investigators include participants who are beginners, intermediates and experts, the expert-performance framework cannot make a priori predictions for the corresponding correlations. However, some studies allow us to report on supplementary analyses that can clarify the nature of the correlation between IQ and domain-specific performance across the attained level of skill.

2.5.1. Chess

In two separate studies Untrainer and his colleagues (Unterrainer et al., 2006, 2011) studied the relation between performance on tests of cognitive ability and chess rating for two samples of 25 and 30 chess players, respectively. The first sample of chess players ranged widely in chess-playing experience ranging from 1 to 40 years ($M = 15.7$) and ranging in skill from 1250 to 2100 ($M = 1683$). The correlation between IQ and chess rating did not reach significance $r(23) = -.076$, $p > .05$ (Unterrainer et al., 2006). The second sample of chess players ranged in chess-playing experience from 5 to 40 years ($M = 22.5$) and ranged in chess ratings from 1209 to 2303 ($M = 1809$, $SD = 273$), and average intelligence test scores for the chess players ($M = 129.7$, $SD = 9.2$). Again the correlation between IQ and chess rating for the chess players was non-significant ($r(28) = -.073$) (Unterrainer et al., 2011). There was substantial variation in chess skill and the standard deviation in IQ scores was comparable to the control groups in each study, thus making the argument that restriction of range was responsible for the non-significant findings less plausible.

Grabner et al. (2006, 2007) collected data on 90 chess players, who ranged considerably in chess ratings from 1311 to 2387 ($M = 1869$, $SD = 247$). An analysis of a subset of 55 chess players, who accepted an invitation to be part of a study collecting EEG revealed a non-significant correlation between general intelligence and chess ratings (Grabner et al., 2006), although detailed results were not reported. In an analysis of the full sample Grabner et al. (2007) reported the average general intelligence ranging from 79 to 141 ($M = 114$, $SD = 14$) and found a significant correlation between IQ and chess rating ($r(88) = .35$, $p < .001$). One issue with this study is that the study recruited participants by advertising that the participants would be provided "information about their intelligence and personality profiles" (Grabner et al., 2007, p. 404). Is it possible that some of the better chess players, who feared scoring poorly on the intelligence tests, might have been less likely to participate? A selective engagement by highly respected chess players with higher IQ scores might have contributed to the observed significant correlation (Dollinger & Leong, 1993). This account is consistent with Unterrainer et al.'s (2011) explanation for the discrepant correlations in their own studies of IQ and performance on Tower of London. Finally, if the researcher had a PhD and was a strong chess players he/she might recruit friends to the chess study with a similar profile, especially if there was no compensation for participating in the study.

2.5.2. GO

Using the national organization for GO in Japan, Masunaga and Horn (2001) were able to recruit 243 amateur players ranging across the whole spectrum of skill and in addition, 20 professional GO players. They also found that none of the 8 traditional intelligence tests correlated significantly with the GO ranking, if we use a Bonferroni corrected p-value of $0.05/8 = 0.006$. The highest correlation was 0.15 and involved backward digit span. The tests of general intelligence also did not significantly differ across the entire spectrum of GO ratings. More recently Jung et al. (2013) compared the IQs of 17 Baduk experts ($M = 93.1$, $SD = 10.1$) to 16 beginners in Baduk ($M = 100.7$, $SD = 12.5$) and the difference was not significant ($p = .062$).

2.5.3. Surgery

Wanzel et al. (2003) reported a significant correlation between spatial ability and surgical performance in beginners (dental students). When they administered the same tests to experienced surgeons (12 surgery residents and 8 staff surgeons), they found no significant correlations. In an interesting study Keehner, Lippa, Montello, Tendick, and Hegarty (2006) studied practice effects on a simulator for laparoscopic surgery and found that "[E]even individuals with relatively poor initial performance could learn the novel perceptual-motor relationships required for the task" (p. 499). Subsequently, Hegarty, Keehner, Khooshabeh, and Montello (2009) have tried to pin down the mediating structure of dental surgical performance and its dependence of domain-specific skills and possibly general spatial skills as a function of the many years of training of dental surgeons. In a recent meta-analysis Uttal et al. (2013) conclude: "Our results clearly indicate that spatial skills are malleable. Even a small amount of training can improve spatial reasoning in both males and females, and children and adults" (p. 370) and recommend that it be trained in a manner similar to reading and mathematics in the regular school system.

2.5.4. Music

Meinz and Hambrick (2010) reported a significant correlation between sight-reading performance and tests of general working memory capacity. Their study included participants with a wide range of technical music skills. For example, they included participants, who had only played an instrument for a year and participants who had taken music lessons for two years but who had never played accompaniments. For their mixed sample of individuals with different levels of music skill Meinz and Hambrick (2010) found that even after controlling for the accumulated amount of piano practice individual differences in working memory significantly predicted sight-reading performance. Their finding are inconsistent with those obtained in studies of musicians with higher skill in sight reading music, such as accompanists (Kopiez & Lee, 2006; Lehmann & Ericsson, 1993). Lehmann and Ericsson (1993, 1996) found the best sight readers were accompanists who were trained to perform with limited or no preparation, and this group of accompanists performed at a superior level on the sight reading tests compared to traditionally trained pianists who typically prepare and memorize their music pieces. For these skilled groups there was no correlation with accumulated hours of music training, but only correlations with measures of the amount of sight reading experience and practice—a pattern completely opposite to that observed for the mixed group studied by Meinz and Hambrick (2010). Consistent with our findings Kopiez and Lee (2006) found that the reported amount of sight reading experience was the best predictor of sight reading performance. Kopiez and Lee's (2006) tests of sight reading were organized by levels of difficulty. When employing the standard significance level of $p = 0.05$ and 2-tailed tests, Kopiez and Lee (2006) only found that working memory reached significance for one level of difficulty, namely Level 3. For the most difficult levels (where acquired skill would be most important) there was no suggestion of an effect of working memory.

Hambrick et al. (2014-this issue) also cite two other studies showing correlations between sight reading and IQ. Luce (1965) found a correlation when they studied a sample with uniformly low levels of sight-reading skill (band members in high school). Salis (1977) studied a mixed group of musicians where roughly half of the participants had no accompanying experience and the other half were either professional accompanists or experienced in accompanying and found a significant correlation ($r(24) = .576$, $p < .01$). In another recent study Hayward and Gromko (2009) found a significant correlation between sight reading performance and spatial ability ($r(68) = .24$, $p < .05$) in a mixed sample of 59 undergraduate students, some of them music majors, and 11 graduate students, who played in ensembles. The results of all these four studies are consistent with a hypothesis that skilled sight reading reflects acquired memory skills that take extensive practice to develop. For a long time sight-reading was viewed as a mysterious skill that was acquired without any direct training. More recently music teachers view this skill as trainable and there are now a large number of research projects studying various methods for rigorous training to improve skill in sight reading (Hagen, Cremaschi, & Himonides, 2012; Mishra, in press).

In a study of general music performance, Hallam (1998) collected data on children playing the violin or the viola, ranging in age from 7 to 16 and with 0–10 years of music playing. Music achievement score was assessed by the grade and level of the high quality examination of the Associated Board of the Royal Schools of Music and this measure was found not be significantly correlated with IQ tests of vocabulary ($r(101) = .11$, $p > .05$) or Raven's matrices ($r(107) = .16$, $p > .05$). However, Hallam (1998) noted that the measure of music performance was highly correlated with length of time learning ($r(101) = .84$, $p < .001$) and estimated total practice time ($r(101) = .67$, $p < .01$).

In sum, our review of the reported correlations between IQ and domain-specific performance found a pattern that is consistent with the predictions of the expert-performance framework based on gradually acquired mechanisms that mediate expert performance.

3. Issue 2: Genetic constrains on attaining expert performance

There are two types of evidence that can show that genetic factors constrain individuals' attainable level of performance in a domain. According to Plomin et al. (2014-this issue) the most compelling evidence involves identification of genes that an expert performer would need to possess as part of their DNA in order to attain the highest levels of performance. Until recently the most prevalent evidence concerns indirect evidence for such genes or combinations of genes derived from estimated non-zero heritabilities for a given level of performance. First I will discuss the evidence on heritability relevant to attaining expert performance.

3.1. Heritability of expert performance

As I wrote in the introduction Plomin et al.'s (2014-this issue) selection of the top 5% of precocious readers in 8th grade does not qualify as adult expert performance for several reasons. The level of reading performance of the precocious readers is attained by the majority of the students at older ages—hence, the absolute level of the precocious 8th graders is not superior to the general adult population. In addition, performing well on comprehension tests designed for 8th graders does not correspond to an expert performance in a real-world domain of expertise. In fact, comprehension of texts appears to depend on the prior knowledge that the particular student has about the domain described by the text (Tarchi, 2010). Finally, precocious reading in 8th grade has not been shown to be predictive of adult expert performance as a writer or academic, as far as I know. This research differs qualitatively from the research described by Wai (2014-this issue), where the relation between precocious performance on SAT-M in 8th grade is directly related to subsequent adult achievement.

Based on our definition of expert performance Plomin et al.'s (2014-this issue) article does not report on a single estimate of heritability for expert performance. Instead their article discusses heritability of performance for the students with the top 5% performance on the test, which Plomin et al. (2014-this issue) call expert readers. By comparing the heritability of reading performance of MZ and DZ twins Plomin et al. (2014-this issue) found that the heritability for "expert readers" is essentially the same as that for the reading performance of the rest of the population of tested 8th graders. Based on this analysis Plomin et al. (2014-this issue) conclude that adult expert performance can be profitably viewed as an extreme end of the normal distribution of performance—a view that I will question below by a review of heritability estimates for the performance of adult individuals, who have attained an expert level of performance.

3.1.1. Why are there so few twins that exhibit expert and elite performance?

Plomin et al. (2014-this issue) mention that "the acquisition of expertise could be due to special environmental and genetic factors that do not affect performance in the normal range" (p. 28) and emergenesis, namely that "rare combinations of many genes are responsible for exceptional performance" (p. 28). The concept of emergenesis was first proposed by Lykken (1982), who proposed that MZ twins were much more similar than would be predicted by prevailing genetic models based on additive effects or dominance relations. MZ twins reared apart displayed similarities not seen with DZ twins reared apart and thus must depend on having the identical configuration of genes, as only MZ twins have. Most relevant for the discussion of the development of expertise, Lykken (1982) proposed that the sudden emergence of genius to seemingly unexceptional parents during the history of our civilization could be explained by emergenesis, namely the genetic offspring would have a unique combination of genes that differed from his/her unexceptional parents. This mechanism explained how "the union of a bricklayer and a peasant woman produced a Karl Friedrich Gauss and also why Gauss's offspring showed virtually none of his mathematical talent" (Lykken, 1982, p. 364). Subsequent researchers have listed many examples of major creative contributors with undistinguished relatives, including Newton, Shakespeare, Michelangelo, Beethoven, and Gauss (Simonton, 1999, p. 131), as well as Faraday, and Benjamin Franklin (Lykken, 1998, p. 30), and Haydn, Schubert, and Schumann (Copp, 1916).

The most obvious method to test the emergenic account of genius (Simonton, 1999) and accounts of expert performance would be to compare the correlation between rates of major creative contributions by MZ twins as compared to DZ twins. Unfortunately, these analyses have not been conducted because Bouchard and Lykken (1999) found very low incidents of attaining the highest levels of performance for even a single member of MZ and DZ twin pairs. The number of exceptional levels of performance was essentially zero for the arts and sciences and a significantly lower number than would be expected by chance. Simonton (1991) examined 2026 eminent scientists and found only one twin (Auguste Picard). Furthermore, there are no MZ and DZ twins among Nobel laureates in the sciences (Simonton, 2008). This striking under-representation of eminent twins, where either one or both members of identical and fraternal twin pairs reach elite levels, led me to conclude that it is virtually impossible to use this type of evidence from twins to support the emergenic account of eminent achievement (Ericsson, Nandagopal, & Roring, 2005). Consequently it will be essentially impossible to evaluate Simonton's (1999, 2005) hypotheses about the critical role of unique combinations of genes by collecting empirical data on twins.

Simonton (2008) even contacted David Lubinski to inquire about the incidents of twins in their studies of the mathematically precocious youths (SMPY) and was told on March 15, 2007 (personal communication) that there were "extremely few twins in the sample" (Simonton, 2008, p. 29). In spite of an extensive search for evidence for data on twins scoring very high in talent searches, I was unable to find any information except one study where only one of the twins in a pair was admitted to the school's gifted programs. Renzulli and McGreevy (1986) found 16 sets of MZ twins and 46 sets of DZ twins meeting these criteria. This ratio of MZ and DZ twins matches the typical ratios in the general population (Kyvik, Green, & Beck-Nielsen, 1995).

The scarcity of exceptional performance among MZ and DZ twins, and the relatively high frequency of precocious MZ and DZ twins in reading during adolescence, leads me to argue that these two phenomena are fundamentally different and thus mediated by different types of mechanisms.

3.2. Identification of unique genes required for attaining expert performance

In the late 1980s Hamilton (1986. p. 64) wrote that "a would-be ballet dancer who has poor turnout from the start probably will never be good, and the attempt to force it can create several knee problems", and Ackerman (2014-this issue) interpreted this statement as clear evidence for the need for a favorable genetic endowment. However, Ackerman (2014-this issue) does not mention that Hamilton (1988, p. 144) two years later wrote that "[M]most of these requirements can only be met by starting early, because many adaptive changes must occur in the skeleton while it is still growing… serious dancers must usually begin their training around the age of 8." Over the following two decades a large body of research shows that intense training influences the levels of turnout, if done before the ages 8 to 11, after which bones are calcified and the hip joint is fixated (for a recent extended review of these issues see Hutchinson, Sachs-Ericsson, & Ericsson, 2013).

In the early 90s many researchers pronounced that in the next years or decades we would be able to identify the particular genes that explained the large estimated variance due to heritable factors. Plomin, Owen, and McGuffin (1994, p. 1733) stated that "the future of behavioral genetics lies in harnessing the power of molecular genetics to identify specific genes for complex behaviors" (p. 1734). In a recent review Plomin (2012, p. 165) concedes that the progress during the last 20 years in identifying the specific genes has been slow: "For instance, initial GWAS [genome-wide association studies] of intelligence, have indicated contributions of many small genetic effects. This is because the genomic differences identified so far between individuals make only a small total contribution to the heritability of this trait—an issue that has been dubbed the missing heritability problem" (p. 165). In a recent review of the progress in identifying individual genes associated with general intelligence (*g*), Chabris et al. (2012) stated that no single gene had been identified in GWAS studies, which would "meet conventional thresholds for significant associations with g (e.g., Butcher, Davis, Craig, & Plomin, 2008; Davies et al., 2011; Seshadri et al., 2007)" (pp. 1321–1322). Chabris et al. (2012) also discuss the problem with reported findings of associations that cannot be consistently replicated, and propose that findings "should be *viewed as tentative* until they have been replicated in multiple large samples" (Chabris et al., 2012, p. 1321, italics added).

Sports is one area where the effects of specific genes would seem especially plausible concerning the highly reliable differences in expert performance in sports, especially sprinting and long-distance running. Everyone is aware of the dramatic domination of the sprinting events by athletes born in USA and the Caribbean Islands and the domination of athletes from Kenya and Ethiopia in long-distance running events. In a review of this evidence in the journal, *Human Genetics*, MacArthur and North (2005) concluded that individual differences in attained elite performance cannot, at least currently, be explained by specific genes. Even seven years later there seem to be a consensus about the failure "to discover a candidate gene that can be conclusively linked to performance" (Tucker & Collins, 2012, p. 557). In a very recent comprehensive review Pitsiladis, Wang, Wolfarth, et al. (2013, p. 553) conclude that "current genetic testing has zero predictive power on talent identification and should not be used by athletes, coaches or parents", but they are hopeful that reliable findings involving genes will be discovered in the future.

How can these claims be reconciled with Plomin et al.'s (2014-this issue) statement about elite performance? "For exceptional athletic performance, a meta-analysis of 366 studies found that a polymorphism in the angiotensin I-converting enzyme gene (*ACE)* is significantly associated with performance in endurance athletes, and a meta-analysis of 88 studies found that a polymorphism in the alpha-actinin-3 gene (*ACTN3*) is associated with power events (Ma et al., 2013)". From my own reading of this paper it seems that the published analysis rejected all articles except for 25 articles for ACE and 23 articles for ACTN3. Ma et al.'s (2013) review also emphasized that they did not consider the effect of performance level and included competitors merely participating in national events. In fact, the study in Ma et al.'s (2013) meta-analysis that showed the greatest differences between sprinters and controls, was conducted by Eynon, Alves, Yamin, et al. (2009, p. 890, italics added), who remarked that the fact that the ACE and ACTN3 alleles: "were *not more prevalent* in our top-level sprinters than in the national-level sprinters implies that this genotype may not be critical to sprint ability, but rather additive"

When we examine the findings for world-class performance the results appear to be clear. Wilber and Pitsiladis (2012, p. 92, italics added) reported in unambiguous terms: "[I] in general, it appears that Kenyan and Ethiopian distance-running success is *not* based on a unique genetic or physiological characteristic." Scott et al. (2010, p. 110) made a similar clear statement with regard to the absent influence of ACE and ACTN3 on world-class sprinters' performance: "The present study has genotyped two of the key candidate genes for human performance in a cohort of the world's most successful sprinters and finds them not to be a significant determinant of their success."

Plomin et al. (2014-this issue) criticizes my proposal (Ericsson, 2007a) for how the engagement in a given activity can modify the expression of genes leading to a different phenotype, which may influence performance. I will therefore offer an example of how two individuals (identical monozygotic (MZ) twins) may develop their physical attributes and abilities differently with different amounts and types of exercise—in spite of the exact same DNA. For example, Williams, Blanche, and Krauss (2005) identified 35 MZ twin pairs discordant for their level of physical exercise. They found that the active twin had a significantly lower Body Mass Index (BMI) and that there was no correlation between the BMIs of the active and non-active identical twins. A subsequent study of 2710 MZ and 2327 dizygotic (DZ) male–male twin pairs McCaffery, Papandonatos, Bond, Lyons, and Wing (2009) found that "[V]vigorous exercise significantly modified the additive genetic component of BMI, which indicated a gene by environment interaction ($p < 0.001$)" (p. 1011). Bogl, Pietilainen, Rissanen, and Kaprio (2009) studied 713 MZ and 698 DZ same-sex twin pairs and wrote in the conclusion: "compelling evidence for the contribution of acquired eating and physical activity patterns on obesity. By using comparative measures within twin pairs, we found that the overall amount of food consumed is the major contributor to obesity, independent of genetic predisposition" (Bogl et al., 2009, p. 538). They asked each twin independently about both their own behavior and also their co-twin's to validate the self-reported information and found that the more obese MZ twin ate more snacks, fatty foods, ate faster, and exercised less.

The phenomenon of obesity can be primarily explained by simply doing more eating and less physical activity than adults with normal weight—no need for acquisition of complex skills though deliberate practice. In contrast, the acquisition of expert performance requires fundamentally different processes and physiological adaptations as the function of the level of the attained performance. Nobody becomes an Olympic level gymnast or World class musicians by simply doing more of the same simple activity that they started out doing as a beginner. In sum, engagement in particular training activities is necessary to create new adaptations at different levels of performance, which implies that heritability estimated for performance of the general population cannot be assumed to generalize to expert performance.

4. Issue 3: Mathematically precocious youths and their adult achievements

The Study of Mathematically Precocious Youth (SMPY) is frequently cited in support of the effects of giftedness on the development of expertise for two major findings. First, individual differences in the extreme end of the ability distribution in 7th and 8th grade is associated with long-term professional success as an adult (Grabner, 2014; Hambrick et al., 2014-this issue; Wai, 2014-this issue). Second, the students scoring very high on SAT-M in 7th and 8th grades can learn mathematics at much more accelerated rates compared to their age level peers (Simonton, 2014-this issue).

Julian Stanley (1977), who founded the project, summarized this work on the mathematically precocious by drawing explicit parallels to coaches in sport. The investigators of SMPY identified high-performing students in 7th and 8th grade, by administering the Scholastic Aptitude Test in Mathematics (SAT-M)—a test normally administered to seniors in high school. Some of these middle-school students performed at or even above the normal performance of seniors in high school. These high-performing middle-school students were given tests of mathematical achievement "to determine what a particular talented student does *not yet know*" (Stanley, 1977, p. 86, italics added), so they would know how much of the mathematics curriculum that the precocious students had already mastered. At this point the youths "are continually helped, facilitated, and encouraged. Each is offered a smorgasbord of educational possibilities" (p. 89). Some students took advantage and skipped grades and were able to graduate college at unusually young ages.

The accelerated courses involved a lot of self-study (Bartkovich & Mezynski, 1981; Fox, 1974). Precocious performance on SAT-M is often associated with high scores also on the verbal portion SAT-V (Benbow & Arjmand, 1990; Brody & Benbow, 1987). Anastasi (1974) hypothesized that in order to successfully study mathematics by themselves from text books individuals need to have developed their verbal abilities.

When we examine the training in music, chess, sports and ballet the pattern of training is very similar to SMPY. For example, when a music teacher meets a new student they would assess the current skill level and design training activities that would provide the next attainable step. The student would then go home to practice some assigned tasks. What makes the training offered by SMPY in mathematics so different from the training in music is that all children attending public and private schools are not given individual training, but are typically instructed together in the class of 20–30 students. The traditional instruction in schools would be more similar to the kind of instruction students would receive during physical education classes or as members of the school band. The SMPY method of providing training that matches students with high mathematics ability to topics that they have not yet mastered most likely insures that their mastery of mathematics is completed in less time and at much younger ages. This finding of speed-up with individualized instruction in mathematics is not in any principled way different from the results demonstrated in the other domains of mastery with individualized instruction, such as music, ballet, chess, and sports.

Within the individual-differences approach the research measuring ability in middle and high school to predict adult achievement is viewed as providing evidence for the claim that "ability matters in the development of expertise in educational and occupational domains, even within the top 1%." (Wai, 2014-this issue, p. 2). Wai (2014-this issue) reports on two data sets to support his claim that individual differences in ability matter, even within the top 1%. The first data set was collect in the project TALENT, where around 400,000 high school students were tested for two days and then invited to respond to follow-up questionnaires 11 years later. Wai (2014-this issue) does not mention that only around 20% responded, which still leaves around 80,000 participants. From the data reported in Table 2 in Wai (2014-this issue) we can estimate the response rates for the top and bottom quartile as the ratio between the reported number and 250 ($0.0025 * 100{,}000$ students in each grade) yielding response rates ranging from 32% (Q4 of general ability in 10th grade) to 47% (Q4 in mathematics ability for 9th grade). It would have been appropriate to report and discuss these relatively low response rates and possible influences of current status on reporting probability. These participants were asked about attending and graduating college, and attending and graduating from doctoral programs. Wai (2014-this issue) only reports that percentages of doctorates among the top quartile and bottom quartile of the top 1% in project Talent differ significantly for mathematics ability but not general ability. There are several selection steps between being a high school student and earning a doctorate. Everyone would agree that getting a BA or BS would be necessary to be accepted by a graduate program. If one were to control for the prestige and academic admissions criteria of the undergraduate college or university that accepted the student, how much variance would be explained in getting accepted to graduate programs and eventually to earning a doctorate? It is important to rule out that academic performance in high school, especially in mathematics, is getting its effect by increasing the likelihood of acceptance in more expensive and prestigious undergraduate institutions, where more professors are actively engaged in research and thus can train students to produce honors theses that will increase the likelihood that they get accepted in graduate school.

The second source of data reported by Wai (2014-this issue) comes from the SMPY project, which identifies students in middle school based on their performance on a regular SAT test, where the top 1% of the students are split into quartiles and the top quartile of students are compared to lowest quartile. These data have similar issues as those with the project Talent, because we know that all universities base their recruitment on SAT or equivalent scores. Consequently, there will be a correlation between a student's SAT score in middle school and their SAT score in high school, which in turn is related to acceptance at the most prestigious research universities, such as Harvard, Stanford, and Princeton. Wai (2014-this issue) does not address how one might be able to control for the effects of the initial acceptance, and the effects of graduation from a more prestigious department and university, which should be quite possible with the data available to the SMPY project.

Wai (2014-this issue) does not discuss the severe methodological problem with the SMPY project data as objective evidence that natural ability is responsible for successfully attaining doctorates and subsequent research and innovation performance. Unlike the project Talent, the SMPY project measured students' performance and then they contacted high-performing students and offered them a wide range of support to help them get appropriate challenging training in mathematics. If mathematically precocious students are offered additional courses and support, how can we know if the resulting benefits are due to the offered support or their superior ability? Swiatek and Benbow (1991) compared the future of precocious students participating in fast-paced mathematics compared to students, who were qualified (equally precocious) but decided not to participate and found that fast-paced participants attended "more prestigious undergraduate colleges" (p. 138) and were "more likely to attend graduate school than nonparticipants" (p. 138). In fact, Wai documented convincingly (Wai, Lubinski, Benbow, & Steiger, 2010) that adult "STEM accomplishments are facilitated by a rich mix of precollegiate STEM educational opportunities that are designed to be intellectually challenging, even for students at precocious developmental levels" (p. 860). According to this paper the SMPY project offered a smorgasbord of benefits including grade skipping and taking fast-paced courses, and the participation in these activities was related to adult STEM success. However, Wai et al. (2010) acknowledges that there was no experimental control over access to these opportunities, and thus the outcomes could be a result of initial ability, family support and resources, motivation and the effect of participation. Benbow, Lubinski, and Suchy (1996, p. 276) states the issue clearly: "the latter cohorts not only received much more assistance from SMPY, they also benefited from the experience gained with the earlier cohorts. Finally, each cohort is successively more able. It is, therefore, difficult to separate out and evaluate these confounding influences". The SMPY project informed the students that they were special (performing in the top 1%) or even "profoundly gifted" (performing in the top 0.01%). The SMPY project "enhanced their confidence in themselves. Yet the most telling finding was perhaps that many students felt they would not have achieved so much academically without SMPY. Many indeed felt that SMPY had enhanced their academic success in lasting and meaningful ways" (Benbow et al., 1996, p. 297). More generally, the labeling of students as "profoundly gifted" is very likely to influence other people's perceptions and expectations. Many of the students that skipped classes in high school were able to graduate early and were often provided early entry to a university. The stature of Julian Stanley at Johns Hopkins University is probably responsible for providing supportive contacts for precocious students enrolling at Johns Hopkins. Swiatek and Benbow (1991) reported that 31/57 or 54% of the students participating in fast-paced mathematics courses attended John Hopkins but only 9/50 or 18% of the nonparticipating students did so. A final concern with SMPY and other programs that offer qualified students more advanced instructional opportunities is there is no experimental control over which students take advantage of the offered training opportunities. It would be reasonable to predict that the adolescents who decided to engage in extra-curricular activities offered by SMPY would differ in important aspects from ability-matched students, who decided not engage in the activities (VanTassel-Baska, 1998). Finally, Coyle et al. (2013) show that the relation between SAT and academic performance is mediated by two different independent factors of roughly equal strength, where only one of them is directly related to general cognitive ability (*g*) and the other one is not. Wai (2014-this issue) does not offer a method for distinguishing the *g*-related influence on academic performance from the other non-*g*-related factor's influence.

The reviewed longitudinal projects are consistent with the expert-performance approach toward developing professional mathematicians, the SMPY project provides the students with training activities that are consistent with deliberate practice and training in other domains of expertise.

5. Issue 4: Predicting performance by the amount of training

In our original paper Ericsson et al. (1993) tried to estimate differences in past practice between groups of expert musicians with different levels of attained performance to test whether higher performance was associated with higher levels of estimated accumulated solitary practice. A significant pattern of differences was found to support the existence of a relation. This analysis is fundamentally different from Hambrick et al.'s (2014-this issue) analyses. They test whether ALL variance in performance can be accounted for by a single estimate of accumulative amount of self-reported practice after correction for lack of reliability in the practice estimates and domain-specific performance.

It is interesting to note that Ackerman (2014-this issue) argued that it was impossible to predict individuals' performance even if we had perfect information of their total amount accumulated practice. For example, he mentions the effects of injuries. Injuries of athletes, musicians, and ballet dancers may fundamentally change their performance after the accident. The injuries may also change what they need to practice to recover a new level of performance. Furthermore, it is fairly common that faulty technique leads to overuse injuries that require complete re-learning of the skills—this type of practice would not be equally beneficial for improved performance. Ackerman (2014-this issue) also mentions the effect of aging. In our original papers we discussed the effects of early experiences and critical periods in several sections, such as "The relation between starting age and performance" (Ericsson et al., 1993, pp. 388–389) and "Early demonstrated abilities assumed to reflect innate talent" (pp. 395–396), and we acknowledge the role of practice at young ages and that some types of practice are only effective or at least far more effective at some ages (cf. turnout in ballet and perfect pitch). Ackerman (2014-this issue) also mentions that practice might be less effective when conducted by older adults than by young adults.

Another factor influencing current performance goes beyond the total accumulated amount of practice and concern the distribution of the practice over the individual's career. Ericsson et al. (1993) discussed the effects of long periods of disuse and how reduced or no practice leads to reductions in physiological adaptations and the need to reacquire these adaptations for individuals who want to reach and surpass their earlier performance. We also discussed how sustained elevated levels of practice were associated with "burn-out" and the need to stop practice for many months (p. 391). I have discussed these effects elsewhere (Ericsson, 1990, 2000), but I want to mention a few findings. For example, Olympic athlete's enlarged hearts reverted back to normal sizes after elimination or reduction of training for a decade (Pelliccia et al., 2002). The performance of older pianists was markedly influenced by their practice during the most recent 10 years (Krampe & Ericsson, 1996). Psychology is replete with reviews of studies showing forgetting and decay of skills with disuse and reduced practice (Rubin & Wenzel, 1996). The effects of all of these factors will not be eliminated by correcting for the reliability estimates of the self-reported practice amounts. If we agree with Ackerman (2014-this issue) that there are all these known additional factors influencing domain-specific performance, then a scientist should argue that Hambrick et al.'s (2014-this issue) stated null hypothesis is false even before any statistical analyses are conducted.

In the original study of the musicians Ericsson et al. (1993) focused on individual differences in past and current practice among highly skilled musicians of roughly the same age, who had steadily increased their weekly engagement in solitary practice (deliberate practice) for the last 10 years, for whom career limiting injuries could not be a factor. Furthermore, we conducted extended pilot interviews with former music students and teachers at the Berlin Academy to identify activities that the musicians viewed as essential to improving their own music performance. Given that I have recently discussed issues related to identifying, collecting data, and analyzing data on estimated engagement in deliberate practice activities I will simply refer to those published papers (Ericsson, 2013a; Ericsson & Moxley, 2012).

Contrary to the hypothesis that superior music performance was due to innate talent and associated with less practice, we found the opposite. The highest performing students had on the average practiced more than less accomplished students in the same program at the music academy. Hambrick et al. (2014-this issue, p. 16) critically note that we "did not report variability statistics for deliberate practice—no standard deviations, variances". Ericsson et al. (1993) did report F-tests so anyone would be able to calculate the pooled within-standard deviation (s_{pooled}), which was 2201 h and the 95% confidence interval around the mean of individual estimates of accumulated solitary practice at age 18 ($M = 7410$) ranged from 2894 to 11,926 h for individual values in the top group (see the calculation of s_{pooled} based on the information explicitly given by Ericsson et al., 1993).[1]

In sum, there is a considerable body of research showing that a high-fidelity assessment of deliberate practice requires sophisticated measurement examining and considering many factors. Measuring self-reported total amounts of practice is a low fidelity approach that can only yield very rough and low-fidelity results.

5.1. Hambrick et al.'s review of self-reported practice

In the meta analysis Hambrick et al. (2014-this issue) included studies meeting two criteria, namely a report with "continuous measurement of practice and cumulative amount of deliberate practice" (p. 7) and a report containing a numerical estimate of a correlation between the two variables. Given the emphasis given by Hambrick et al. (2014-this issue) as well as Ackerman (2014-this issue) on the wide range of estimated accumulated deliberate practice to attain the level of chess master I think that it is interesting to note that similarly extreme values were found by Ericsson et al. (1993) in their original study published over 20 years ago. This leads me to compare the methods used by these two studies. Ericsson et al. (1993) developed categories of practice activities carefully and had in depth individual interviews to extract reference points, such as years of changing music teachers, followed by a year by year estimation with the help of an experimenter along with ratings of different practice activities and collection of daily diaries for a week to validate the categories and estimates.

In direct contrast, Campitelli and Gobet's (2008) procedure involved collecting anonymous questionnaires, where "the questionnaire was left visible at the reception desk in the Círculo de Ajedrez Torre Blanca, one of the most important chess clubs in Buenos Aires (Argentina)" (p. 447). As part of the questionnaire there was one page with a matrix with age on one axis and two types of practice, namely "studying chess alone at each age" and "studying or practising with other chess players, including tournament games" (Campitelli & Gobet,

[1] The formula for the F-test for differences between groups is $F = (M_1 - M_2)^2 / (MS_w * (1/N_1 + 1/N_2))$ and the $F(1,27) = 4.59$ for the difference between the top group of violinists ($N_1 = 10$, $M_1 = 7{,}410$) and the middle group ($N_2 = 10$, $M_2 = 5301$) data reported by Ericsson et al. (1993, p. 379). When solving for MS_w the result is 4,845,186.27 or $S_{pooled} = MS_w^{0.5} = 2201$. Solving for MS_w for the second F-test gives the same result within rounding errors.

2008, p. 448). By multiplying these practice estimates by 52 weeks and adding them all up Campitelli and Gobet's (2008) computed their estimates of life-time practice. Data was collected from 104 respondents, but Campitelli and Gobet's (2008) only analyzed 90 participants and did not describe the objective reasons for discarding 14 of the collected questionnaires.

Gobet and Campitelli (2007) were not able to collect test–retest reliability of the information nor other independent data supporting the reliability and validity of the collected data. Given that most of their data was given anonymously, retesting participants would be impossible. Even more problematic with anonymous questionnaires is the fact that it is not possible to further interview Gobet and Campitelli's (2007) chess master who reported practicing 728 h of individual practice[2] and 1612 group practice (2340 total hours of practice). Does Hambrick et al. (2014-this issue) claim that the reported amount of practice is perfectly accurate—without any error? Typically any data point is reported with a confidence interval based on reliability and validity estimates for the error variability, but these were never collected in this study. To treat this single value as factual appears to me to be highly questionable from a scientific point of view. More generally it would be very interesting to interview this individual to assess more details about the nature of the individual practice and a detailed account of tournament outcomes and the increase of chess rating. If this individual is anonymous then this type of detailed study and validation would not be possible. If on the other hand, one is simply looking for the lowest estimates of chess study it is interesting to note that Howard (2012) in his internet study included "at least one elite chess player that had never studied chess (average of zero hours of study) and another elite player who had studied an average of 60 h per week across his or her entire chess career (which would mean that he or she studied nine hours every day since he or she initiated serious study many years earlier)" (Ericsson & Moxley, 2012, pp. 651–652). It is surprising that Hambrick et al. (2014-this issue) did not cite Howard's (2012) data for evidence of an elite chess player, who had *never* studied chess—an even more compelling exception to the idea that chess study is required for expert performance.

A more fundamental problem is highlighted by the chess master with the highest amount of self-reported individual practice. Gobet and Campitelli (2007) identified a chess master, who reported having engaged in a total of 24,284 h of individual practice.[3] According to Gobet and Campitelli's (2007) procedure section they only asked about the chess players' current chess rating—not their highest rating and what year they attained that rating. It would be nice to have Gobet and Campitelli (2007) conduct a re-analysis that would identify the amount of practice required prior to first attaining the rating of master, but the anonymous questionnaires may not allow gaining such additional information. Their current report does not rule out the possibility that the chess master in question attained the level of chess master many years earlier, and thus his/her accumulated estimate would include chess training for those years after he/she first attained the chess rating of the master. Krampe and Ericsson (1996) identified a similar problem in their study of both a young (around 20–30 years old) and an older group (around 50–60 years old) of expert musicians. The older experts had accumulated over 50,000 h of solitary practice, which was well over twice as much as younger expert pianists with a similar level of current performance. This illustrates the problem with using a single estimate of accumulated practice. Krampe and Ericsson (1996) calculated several different estimates that better characterized old and young music experts. One of the estimates calculated the amount of practice until age 20 (an estimate calculated in Ericsson et al. (1993)). Another estimate that was predictive of different aspects of music performance for the older musicians which was the number of hours accumulated during the last 10 years.

More recent studies show that it is not the total number of hours of practice that matters, but a particular type of practice that predicts the difference between elite and sub-elite athletes. For example, elite middle distance runners do not train more hours than sub-elite runners, but rather they engage in more hours of interval training (Young & Salmela, 2010). Only by focusing on the development of particular adaptations and mental mechanisms (see Fig. 1) can we identify those particular training activities that are most effective (see Ericsson, 2013a, for a more extended discussion of this point).

6. Additional issues raised by contributors of the special issue

One of the remaining issues concerns how children and adolescents attain precocious performance in mathematics in 7th and 8th grade (Wai, 2014-this issue) and how prodigies attain their superior performance at very young ages (Ruthsatz et al., 2014-this issue). Another issue concerns the frequent references to exceptional individuals, who attain world-class performance within very short periods of time, and thus would suggest abnormally rapid learning and skill acquisition (see Ackerman, 2014). Finally I want to discuss Wai's (2014-this issue) evidence on successful individuals in our society, such as member of congress, billionaires, federal judges, and CEOs of Fortune 500 companies, and their inferred IQ based on attending elite colleges.

6.1. Necessary conditions for the development of precocious performance

Nobody would seriously argue that precocious children and adolescents discover more advanced mathematics completely by themselves. Bartkovich and Mezynski (1981) showed that precocious children had not encountered their advanced knowledge of mathematics in school. More recently Campbell (1996) found that mathematically precocious students had acquired the knowledge by self-study using popular books. One of the most extensive studies of how children with gifted IQs grow up was conducted by Gottfried, Gottfried, Bathurst, and Guerin (1994). In their longitudinal study of a large group

[2] Surprisingly, Hambrick et al. (2014-this issue) reports the lowest value for a chess master as 832 h instead of the 728 h as reported by Gobet and Campitelli (2007, p. 166) without providing an explanation for the difference.

[3] Surprisingly, Hambrick et al. (2014-this issue) reports the highest value for a chess master as 24,284 h instead of the 16,120 h reported by Gobet and Campitelli (2007, p. 166) without providing an explanation for the difference.

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

of children they contrasted the characteristics of those children that later received scores in the gifted range with children receiving lower scores. They found that the gifted children's environment and daily activities involved more variety of stimulation even as early as the first and second year of life. From age 3 they had more toys, games and materials, and were more stimulated by language and academic activities. They particularly stressed the finding that the gifted children requested more activities than the non-gifted children at age 8. Gottfried et al. (1994) even found "[E]evidence for gifted behavior in non-gifted, but environmental advantages were not present, parents did not recognize and nurture the accelerated development, or the child's own motivation was insufficient" (p. 171). More generally, the gifted children and adolescents were rated as being "significantly more goal directed; having greater absorption in tasks as indicated by longer attention spans" (p. 172). The gifted children had greater environmental resources (higher socio-economic status) and were "predominantly firstborns, and quite often only borns" (p. 174).

The research focusing on the early development and family environment of students with very high mathematical performance has been studied within the context of Math Olympians (Campbell, 1996). In 1995, 350,000 students took the same exam in mathematics—the American High Mathematical Examination. The top 15,000–20,000 scores were invited to take a second exam, where 140 top scorers were selected for a third exam and finally the top six scorers (plus two alternates) were selected to attend a 4-week training camp to represent the USA in the international Math Olympiad—this group is 0.002% of the original population of students. Campbell (1996) sent questionnaires to the Olympians and their parents soliciting information about the family background of this elite group. The students were primarily first-born in small families, who had professional parents with a high SES with books and many other resources—roughly equivalent to the eminent scientists studied by Anne Roe (1953a, 1953b). The Olympians had "a strong belief in developing a self-taught orientation. Many of them reported learning much of their math "on-their-own" from books" (Campbell, 1996, p. 514). They and their parents emphasized "the importance of having numerous books available in their homes" (p. 514) including popular math books. Subsequent research findings on Olympians in Mathematics are consistent with those of the longitudinal study by Gottfried et al. (1994) where social family factors, such as high socioeconomic status and effort attributions, were important predictors of becoming an Olympian in Mathematics in the USA and Finland (Nokelainen, Tirri, & Campbell, 2004), and in Taiwan (Wu, 1996). Of particular interest is the research on individual differences within the highly selected group of Olympians in Mathematics, Chemistry, and Physics from USA, Germany, and Finland in terms of their adult success. Nokelainen, Tirri, Campbell, and Walberg (2007) found differences in the childhood experiences of the Olympians, who became the most academically productive third as adults (over 1500 publications, patents, and software products) and the least productive third of the Olympians (less than 20 publications, patents, and software products). The most productive third reported having experienced a more supporting family environment. Only future research will help us identify if there are environmental variables that can promote the development of students' performance that will prepare them for a productive adult professional career.

There is some very interesting research on gifted students that has employed the methodology of the expert-performance approach to attempt to assess how gifted students' thinking compares to non-gifted students. These studies have collected "think-aloud" protocols from both gifted and average students, while respond to the same tasks and solving the same problems to assess differences in the strategies used by the two types of participants. When participants of the same chronological age are compared, gifted students display better performance and use superior strategies in reading texts (Fehrenbach, 1991) and in learning with hypermedia (Greene, Moos, Azevedo, & Winters, 2008). Of particular interest are studies, where investigators have matched younger gifted children with older children with average performance. When gifted 9-year-old children were compared to average 13-year-old adolescents they "were found to be very similar, as was the frequency of different responses to the questions, suggesting that many of the mathematically "gifted" are not qualitatively different in the problem-solving approach from students of average ability, but are merely precocious" (Threlfall & Hargreaves, 2008, p. 83). Further research on gifted students using methods from the expert-performance approach would seem to have the potential for creating genuine bridges between the two approaches. We need the kind of careful studies of gifted children's development of their performance, as researchers who are studying the development of vocabulary and language have conducted including daily diaries of new developments (Clark, 2009).

6.2. Case reports of individuals attaining expert performance rapidly

The original claims for innate talents were based on achievements that seemed inexplicable in terms of the normal gradual acquisition of performance. In an influential book Gardner (1973, 1994) explained his belief in the fundamental and virtually exclusive role of innate genetic factors: "Further evidence of the strong hereditary basis of musical talent comes from a number of sources. Most outstanding musicians are discovered at an early age, usually before 6 and often as early as 2 or 3, even in households where relatively little music is heard. Individual differences are tremendous among children, and training seems to have comparatively little effect in reducing these differences." (p. 188). Many researchers (Howe, Davidson, & Sloboda, 1998; Treffert, 1989) and including myself have discussed this evidence and proposed alternative skill-based explanations in terms of early interactions between parents and the gifted child based on the methods for training any child in music according to the principles of Suzuki (Ericsson, 2002; Ericsson, Roring, & Nandagopal, 2007; Lehmann & Ericsson, 1998). There are a number of very popular anecdotes about precocious performance, but much of that evidence cannot even be verified by objective sources. For example, the popular childhood anecdotes about Gauss' mathematical genius were first reported by Gauss himself as an old man. These and other anecdotal reports nearly always lack independent verification and are therefore not even considered in modern biographies of Gauss (Bühler, 1981). In contrast to such unverified cases, Ackerman (2014-this issue) reported on objective performances that were attained during public competitions. He mentions three athletes that reached their level of international performance in a very short time compared to the data on most athletes.

When the superior performance is demonstrated in public, the only remaining question concerns the quantity and quality of training that preceded the performance nearly always engaged in isolation or a training facility. For example, Ericsson and Faivre (1988) discussed evidence on the development of mental calculators (ability to mentally multiply large numbers without any external memory aids) and proposed that young shepherds who had to spend long periods including nights by themselves with their sheep, would discover an attention-demanding activity to reduce their anxiety and fears. The shepherds would be counting sheep most of the time so the step to increasing the complexity of this activity to mental multiplications would be a simple step. To maintain the need for full concentration they would then increase the size of the numbers multiplied. Mitchell (1907) discusses how these observations can account for why most of the famous calculators known to him had grown up being shepherds, where their skill development occurred in isolation prior to its first demonstration in public.

Another issue involved in measuring speed of acquiring a skill concerns the determination of the start of the training. For example, how soon a scientist is able to produce major discoveries and theories depends on what age is used to identify the start of scientific research (Holmes, 1996). One method would be to use the date of their successful defense of the doctoral dissertation. The dissertation research for some scientists is later hailed as a major research contribution, and according to this method of selecting the start age these scientists would be viewed as requiring no training time before making a major contribution to science. If we on the other hand start searching for time that all the knowledge and skills necessary for making the discovery were attained, such as math and science in middle and high school, we would arrive at a very different estimate for how long it would take to prepare someone to be able to generate deliberately a new major research contribution.

A more careful examination of the three examples cited by Ackerman (2014-this issue) shows a similar problem. Two of the three athletes had started training as athletes with many years or even decades of intense training before they started with the particular sport. For example, Glover was an accomplished athlete and won an athletic scholarship in 2002, which was well before she started rowing training in 2008 and her gold medal in 2012 (Glover, 2013). Similarly, Chrissie Wellington says in an interview that she engaged in sports as a kid, and she started to run regularly during her work on her master's degree, well before she started training for the triathlon (Wellington, 2013). Finally, Donald Thomas had trained in sports for a very long period and had even competed in high jump in high school in the Bahamas according an article in *New York Times* (Cleary, 2007). Other cases of rapid success after starting with the specific sport concern individuals with a long history of sport involvement and nearly always concern sports, where the technical skills are less, such as swimming, running, and bicycling (Johnson, Tenenbaum, & Edmonds, 2006). The level of discussion of these individuals with reports on exceptional rapid development of international performance would be so much more meaningful if the researchers took the time to describe and study the longitudinal process of development from the first engagement in relevant activities. Even collecting retrospective reports (Côté, Ericsson, & Law, 2005) of all relevant practice activities from early childhood along with performance information and training data on aspects, such as maximum weight-lifting and running times for practice runs, from the initial start of the new sport would allow us to identify general effects of training that can transfer, as well as collecting clearer evidence on the nature of potential innate advantages.

6.3. The difference between socially recognized experts and expert performers

The expert-performance framework restricts its research to objectively measurable performance. It rejects research based on supervisor ratings (Schmidt & Hunter, 2004) and other social indicators, because there is no direct way to identify if these ratings correspond to measureable differences in representative performance and how these performance differences can by objectively and reproducibly measured.

Consistent with the acceptance of social ratings and indicators by the individual-differences approach, Wai (2014) reports on an analysis of socially recognized individuals, such as the top Federal judges, billionaires, CEOs of Fortune 500 companies, senators, and members of the House of Representatives (Wai, 2013). Wai (2013, p. 203) claimed that "America's elite are largely drawn from the intellectually gifted, with many in the top 1% of ability". By examining the prestige of the universities that these individuals attended as undergraduates and the attained level of education, such as doctorates, Wai (2013) infers that these individuals must have been academically very successful and thus likely part of the top 1% of the intellectually gifted. Wai (2013) admitted that his analyses are based on average scores for SAT and ACT for recent years, and thus does not address whether individuals were accepted based on other, most likely lower, criteria, such as legacy (parents attending the college) and/or donations to the college. He did not discuss the issue that the mean age of his cohorts was 56 for the youngest group (members of the House of Representatives) and 66 for the oldest (the billionaires) with the oldest members being 97 years old. It is likely that the methods for accepting students were different 40–60 years ago, when many of the individuals were accepted as undergraduates at their respective universities.

Wai (2013) did not collect information on the SES of the parents, the family structure (see earlier discussion of first-born and small families), attending preparatory private schools, and financial resources to attend the more expensive prestigious colleges and graduate schools to show that these factors could not explain the higher percentage of "successful" individuals attending elite colleges. With knowledge of the particular college attended Wei might be able to assess the effect of attending a particular college, such as Harvard, when compared to other less famous colleges with comparable criteria for SAT scores for admitting students. I am unable to see how the data presented by Wai (2013) compels anyone to believe that innate ability reflected by

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

the performance on a SAT test influenced the success of those with the highest SAT scores.

More significantly for the expert-performance approach, Wai (2013) does not discuss if and how these individuals have demonstrated reproducibly superior performance on some task, as required by the expert-performance framework. I would like to know, what it is that Senators and Members of the House of Representatives are able to do that their less successful peers could not do. For an incumbent to keep winning re-elections in safe districts does not qualify—at least with our criteria for reproducibly superior performance for standardized tasks. Even more problematic is the fact that most CEOs, billionaires, and politicians have large teams of hired help, who must have been responsible for significant aspects of the famous person's successful attainment of wealth, power, and influence. Consequently, we cannot attribute the success of these individuals to their unique individual performance.

The performance difference between being socially recognized as an expert and being able to perform at a reproducibly superior level of performance is now well documented. For example, when individuals, based on their extensive experience and reputation, are nominated by their peers as experts, their actual performance is often found to be unexceptional in domains of expertise in medicine (Ericsson, 2007b), auditing (Bédard & Chi, 1993), language translation (Jääskeläinen, 2010) and nursing (Ericsson, Whyte, & Ward, 2007). There is also very limited evidence of general improvements in performance as the result of extended professional experience. In fact, in some cases during the time since graduation from medical or nursing school, there is even evidence for decrements in performance, most likely due to forgetting (Choudhry, Fletcher, & Soumerai, 2005; Ericsson, 2004, 2007b). A recent meta-analysis of clinical reasoning showed a similar effect—with a noticeable disadvantage for people without any experience, but little if any advantage for additional experience beyond the initial exposures (Spengler et al., 2009). Similarly, Philip Tetlock (2005) compared predictions from hundreds of experts in different fields with well-informed, non-experts and was able to dispel the myth that experts' forecasts are superior. Very relevant to Wai's (2014) argument that superior education of successful individuals mediates their success, there are at least some domains, such as psychotherapy (Montgomery, Kunik, Wilson, Stanley, & Weiss, 2010) and teaching in K-12 (Hanushek & Rivkin, 2010), where the level of training beyond the required minimum, such as additional master and doctoral degrees, has no measureable favorable impact on the ability to improve psychological clients' health or help students learn better.

One also needs to be aware that supervisors' ratings of performance may be correlated with a particular persons' IQ. but not with measures of their objective performance. Cole and Cole (1973) found a non-significant relation between IQ and publication quantity and quality, but a significant relation between IQ and prestige of the individuals' department. They suggested that IQ might influence hiring decisions independent of measures of scientific productivity. Similarly, Vinchur, Schippmann, Switzer, and Roth (1998) found that IQ was not correlated with objective sales performance ($r(1308) = .02$, $p > .05$) and the attenuated correlation was 0.04 after correction for criterion unreliability and range restriction, but significantly correlated with supervisor ratings of sales performance ($r(1229) = .23$, $p < 0.001$) and the attenuated correlation was .40 after correction for criterion unreliability and range restriction. The scarcity of objective valid measures of experts' performance along with demonstrations of a dissociation between supervisor ratings and actual performance lead the expert-performance approach to disregard findings based only on social criteria and ratings.

Finally, it may be tempting and convenient to rely on self-reported ability and success. Hambrick et al. (2014-this issue) cited a study by Vinkhuyzen, van der Sluis, Posthuma, and Boomsma (2009), where they had twins rate their level of skill for a number of different domains, such as arts "Arts referred to artistic and creative activities (painting, acting)" (p. 382) and chess "Chess referred to the ability to play games like chess, backgammon, and mah-jong" (p. 382). This study stands in clear isolation as I don't know of any other studies that have collected data only on self-ratings of ability.

6.4. Do general personality traits predict expert performance?

In their articles Simonton (2014-this issue) and Hambrick et al. (2014-this issue) argue that individual differences in general personality traits are likely important predictors for attained expert performance. Unfortunately, neither of them cites studies that unambiguously demonstrate evidence that general personality traits are causally related to the acquisition of reproducibly superior performance.

The expert-performance approach is focused on explaining the structure and acquisition of expert performance and has found that deliberate practice is the most promising proximal variable with a plausible mechanism for explaining change (improvement) of performance. As our knowledge of how deliberate practice is initiated, maintained, and causes different changes in mechanisms mediating performance at that time, it would be productive to ask questions about which variables might influence the engagement in practice.

Our current knowledge about general personality traits has not attained a state where it can be productively related to expert performance. I have developed this argument in a full-length chapter (Ericsson, in press) so I will only summarize my argument. Expert and elite performers start engaging in domain-related activities typically at a very young age. It is therefore possible that the act of engaging effectively in the domain with daily deliberate practice influences their preferences and reactions to items on personality tests—thus having the engagement in the domain influences the preferences for activities rather than having the endorsement of general personality traits influence their selective engagement in activities. Especially if scientists and artists are tested only as adults, any differences in the responses in the questionnaires measuring personality could be the result of decades of adaptation to their engagement in deliberate-practice activities in the arts and science during childhood, adolescence and early adulthood.

In a recent review Feist (2006) acknowledges that there may be bi-directional effects between success in science and questionnaire responses related to general personality traits. He acknowledges that there have only been two longitudinal studies of personality in scientists and both were conducted after the individuals had already entered their scientific careers. In one of them Feist and Barron (2003) compared tests administered at age 27 with tests given at age 72 years

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

old, and found that some of the personality traits had changed significantly while pursuing a scientific career. Predictions for how personality is likely to influence choice and persistence in professional careers are likely to have their effects at the start of the individuals' careers—presumably in childhood and adolescence and early adulthood. Until there are studies collecting data for these developmental periods, I think that it will be impossible to distinguish whether the choice of career, with its intense engagement in practice, influences responses to personality test items or traits influence responses to personality items and cause individuals to pursue certain careers successfully.

More generally, it will be necessary to empirically support claims about correlations between performance and personality, heritabilities of personality traits and other characteristics of general personality traits by data for the target population of expert performers, rather than based on studies of the general population. We need to be very clear about what we know and have already demonstrated empirically in samples of expert performers.

7. Concluding remarks on creating bridges between the two approaches

My involvement in the discussion about evidence on innate limits for attaining expert performance started with my research with Bill Chase, where we found that an average college student could surpass every validated feat for memorizing spoken digits after 200–400 h of training. In our original paper we (Ericsson et al., 1993) questioned Galton's (1869, 1979) influential hypothesis that immutable innate factors set limits for how much an individual would improve with practice thus attributing exceptional performance to individuals born with innate talent. I totally agree with Plomin et al.'s (2014-this issue) distinction between individual differences in performance as they are observed right now "what is" from the attainable performance that individuals could attain with more effective practice activities "what could be".

In an early response to Ericsson and Charness (1994) article in *American Psychologist*, Howard Gardner (1995) conceded that he did not know of any objective data for innate talent but argued that his "response relies on considerations of logic and common sense, whereas their [Neil Charness and my] conclusions rely on 'data'" (p. 803). A similar attitude is displayed by Hambrick et al. (2014-this issue), who even today dismiss our arguments on a similar common sense basis and wrote that "Gardner (1995) commented that the deliberate practice view 'requires a blindness to ordinary experience'"(p. 6). The view that no rigorous evidence is necessary to dismiss our ideas, suggests that any type of evidence is even better. This attitude might explain why some of the contributors to the special issue are reporting contaminated evidence, anonymous questionnaire data without any assessed reliability, self-reported data on the participants' abilities and performance, and the rejection of a null hypothesis shown to be incorrect even prior to any data collection. It would be very helpful if the contributors in their responses to my reply would help us clarify what empirical evidence that they accept as the best currently available evidence that innate talent constrains the level of expert performance that can be attained with deliberate practice.

In the last decades there has been a massive accumulation of detailed knowledge of the complex brain and physiological mechanisms mediating reproducibly superior performance in a wide range of domains (Ericsson, Charness, Feltovich, & Hoffman, 2006). Any serious proposals for explaining the emergence of the complex mechanisms of expert performance cannot merely state that it is innately programmed or acquired through training. My own research has focused on *how*—the type practice that can predictably improve performance, especially at higher levels. The Nature–Nurture dichotomy is no longer scientifically meaningful and there should be a "replacement of the questions, "Which one?" and How much?" by the more basic appropriate question, "How?" (Anastasi, 1958, p. 206). There are several researchers of intelligence, who have been interested in the issue of how general intelligence and acquired expertise mediate skilled and in particular expert performance. It is particularly interesting to note that they are some of the most respected researchers in the individual difference tradition, who have on various occasions conducted research on reproducibly superior expert performance. During the Buros-Nebraska Symposium on Measurement & Testing in 1985 I was fortunate to spend some time talking to Arthur Jensen. I was impressed by his strong commitment to finding a plausible neural mechanism for individual differences in general intelligence, *g*. In his talk and the resulting chapter Jensen (1987) criticized earlier proposals by Spearman and Burt to provide a mechanism for the effects of *g*. Jensen (1987) had explored the possibility that *g* reflected the speed of elementary information processes, but argued that "the *g* factor reflects the more fundamental attribute of mental speed" (Jensen, 1987, p. 122). During one of our conversations he mentioned that he had been able to study one of the most remarkable mental calculators in the world, Shakuntala Devi of India. Among many feats, she was able to multiply a 9-digit number with 4-digit number in less than 20s during an interview with New York Times. He told me how excited he was about testing her and "whether such exceptional performance depends on the speed of elementary information processes" (Jensen, 1990, p. 259). He told me that he was so surprised to find that in spite of Devi's amazing speed on mental calculation tasks, her performance on the Raven's matrices "unexceptional" (Jensen, 1990, p. 266), her backward digit-span was 4 digits. Most surprisingly, her reaction times and other psychometric characteristics fell in the normal range. When Jensen (1990) later published his research he cited my earlier work with Bill Chase summarized in Ericsson (1987, 1988) as evidence that "with great amounts of practice high levels of expertise in various skills can be attained by quite ordinary people" (p. 272). He concludes that "[S]some kind of motivational factor that sustains enormous and prolonged interest and practice in a particular skill *probably plays a larger part* in extremely exceptional performance than does psychometric g or the speed of elementary information processes" (p. 259, added). As far as I can tell Jensen never conducted any further studies of expert italics performance.

Ackerman (2014-this issue) is arguably the world's authority on the relation of intelligence and expertise. In a recent chapter on this topic Ackerman (2013) concluded his chapter: "Whether higher intellectual abilities are *necessary* for acquisition of such levels of expertise *is not directly known*, because gatekeepers to entry for these occupations depend on intellectual ability tests for selection in the educational or occupational programs" (Ackerman, 2013, p. 857, italics added). He re-iterates his earlier expressed view (Ackerman, 1987) regarding closed tasks, such as execution of procedures in high predictable environment, "the influence of intellectual abilities diminishes with increasing practice, as motivation, effort and persistence increase in influence" (Ackerman, 2013, p. 857). Consistent to my own review Ackerman's (2013) finds no current reproducible empirical evidence supporting a significant relation between expert performance and intelligence among skilled and expert performers. It is, also, noteworthy that Ackerman (2014) did not cite many types of evidence cited by other contributors to this special issue—for reasons, I hope, might be similar to those that I have described in my response.

In this response I have tried to discuss all the evidence that Ackerman (2014) cited in support of the necessity of innate superiority in general cognitive ability to attain expert levels of performance, except, at least one. Ackerman (2014) stated the following cognitive limitation, namely: "[T]there has not been a single study that has demonstrated the attainment of expert memory among severely, moderately, or even borderline intellectually retarded subjects". Recalling our stated criterion that the individual needs to be able to engage in deliberate practice, it might still be possible to find individuals with IQs below 85 (the upper bound for borderline intellectually disabled), who have reached an expert level in a domain involving memory. I have already reviewed the evidence for the existence of such individuals in chess and GO. The challenge is to find a memory domain, where regular people would be motivated to acquire this type of expert memory skills with real-world monetary incentives. One of the few examples that I have been able to find concerns the need to pass the test of knowledge about some 25,000 streets and what the best routes are between many different pairs of points to become a taxi driver in London. Wollet and Magure (2011) tested taxi drivers, who had passed the tests and those that had not, in terms of their intelligence. They found that taxi drivers, who successfully passed the test had a verbal IQ ($N = 39$, $M = 97.7$, $SD = 6.3$) and a Matrix score ($N = 39$, $M = 11.9$, $SD = 2.1$) and the drivers who did not pass the test had virtually identical means for verbal IQ ($N = 20$, $M = 98.7$, $SD = 3.5$) and the Matrix test ($N = 20$, $M = 12.2$, $SD = 1.98$). Woollett and Maguire (2011) found that successful taxi drivers had greater changes in their brains, but conceded that "The trainees that qualified [successfully passed the tests] may have had genetic predisposition toward plasticity that the nonqualified individuals lacked" (Woollett & Maguire, 2011, p. 2113). Hopefully, Ackerman (2014) would be willing to specify the evidence for the existence of particular unmodifiable limits on attainable performance in spite of an ability to engage in deliberate practice to help our discipline determine exactly "what could never be".

Ackerman (2014) dismissed our research on expert performance as a flawed application of the psychometric approach with "(1) small samples, (2) restriction in range, (3) either poor or otherwise limited measures of traits" (p. 24) and argued that the appropriate measurement of multiple traits requires many measures and that this requirement "has never been met by the kinds of investigations that seek to determine individual differences determinants of expert/elite performance" (p. 28). I would propose that there is at least one study that meets that requirement, but was not cited by Ackerman (2014). It seemed to me that the study by Masunaga and Horn (2001) meets these criteria. John Horn, who is generally recognized as one of the major contributors to research on the structure of intelligence, conducted a study of GO-players. In this study they collect data on performance on 263 GO-players with 62 beginners, 89 intermediates, 92 advanced players and 20 professional GO-players. They collected data on 4 major psychometric factors of cognitive ability, where each factor was measured with two parallel tests and they also "translated" the abstract psychometric tasks into 5 matching tasks that were meaningful within the domain of GO (Masunaga & Horn, 2001)—a total of 13 psychometric tests. For example, one of the reasoning tasks instructed participants to select one of five figures that had a dot in the same location as the presented target figure. The corresponding GO reasoning task involved selecting the best next move for a GO position. A traditional short-term memory task asked participants to reproduce as many presented squares with different side markings after a minute of study. The matching memory task in GO asked participants to reproduce the location of stones in GO position after a brief presentation. In the discussion of their paper Masunaga and Horn (2001, p. 308) concluded that the performance on the 4 psychometric factors was consistent with earlier research. Performance on the tasks translated into the GO domain measured reliable abilities that correlated highly with performance ranking in GO. Most importantly to the issue of general cognitive abilities, these expertise abilities were "found to be *largely independent* of the inductive reasoning (Gf), STWM, and Gs abilities that heretofore had been regarded as capturing the essence of the reasoning, memory, and speed components of human intelligence. These results suggest that there is more to human intelligence, namely expertise abilities, than has been measured in traditional IQ tests, and that this "more" may be found at high levels in adulthood" (Masunaga & Horn, 2001, p. 308). In a subsequent chapter Horn and Masunaga (2006) proposed "A Merging Theory of Expertise and Intelligence", where they propose how to integrate "the extended theory of fluid (Gf) and crystalized (Gc) intelligence" (p. 588) with an expertise theory, "where the main ideas of this theory are well described in Ericsson (1996), Ericsson and Charness (1994, Ericsson and Kintsch (1995), Ericsson and Lehmann (1996) and Ericsson, Chapter 38 (Ericsson, 2006b)" (p. 588). Consistent with their view that expertise abilities measure independent abilities from the traditional measures of intelligence, Horn and Masunaga (2006, p. 605) conclude that whereas fluid general intelligence "decline with age in adulthood at all levels of expertise" expert reasoning and memory relevant to GO "increase with level of expertise, and to the extent that there is deliberate, well-structured practice to develop and maintain expertise, these abilities increase with age in adulthood. These abilities exemplify more nearly than the others the full capacity of human intelligence" (Horn & Masunaga, 2006, pp. 605–606). With his unexpected passing in 2006, Jack Horn cannot to be part of the discussion today, but I would like to cite Horn and McArdle (2007 p. 242) in their subsequently published handbook chapter: "Extended Gf–Gc theory does not adequately describe abilities that appear to be quintessential expressions of human intelligence—in particular, abilities that reach their

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

peaks of development in adulthood." I can only wish that empirical and theoretical work by (Horn & Masunaga, 2006; Masunaga & Horn, 2001) would receive the attention that it deserves. Their study tested middle-aged participants with an average age of 55 years of age, but I cannot see anything about this study by one of major contributors to the structure of intelligence that would make its conclusions irrelevant to the study of expertise and expert performance.

I find Horn and Masunaga's (2006) proposal for two types of ability factors (one domain-specific and the other general) as a really valuable contribution to the study of expert performance and general cognitive ability, and perhaps the closest to a useful expression of our ideas of expert performance within the individual-differences framework.

After my review of the contribution to the special issue on "the Development of Expertise" I am getting increasingly convinced that the expert-performance framework and its case-based methods offer a superior approach to the study of expert performance than the individual differences framework. Any method, like the individual difference approach that requires large samples of individuals to identify general traits to account for individual performance will never be able to account for the very highest levels of performance—a level of performance attained by less than handful individuals. It will similarly have difficulty describing the changes in traits responsible for the historical changes in the level of performance over the last century, in domains of expertise, such as dance, chess, and sports (Ericsson, 2006a, 2006b; Ericsson & Lehmann, 1996; Ericsson et al., 1993) The expert performance approach was explicitly designed for studying complex performance with individual cases and small samples (Ericsson, 2006b) to first describe the acquired structures mediating performance including differences between individual expert performers. It provides methods for tailoring experiments to experimentally test hypotheses of the structure acquired by a single participant. To make its results more relevant to the individual differences approach it will be important to encourage scientists to search for generalizable aspects of the acquired cognitive structures. This is very similar to the induction of similarities from detailed validated descriptions of expert memory performance that Bill Chase, Walter Kintsch and I did to develop skilled memory and its extension to LTWM (Ericsson & Kintsch, 1995). Anybody interested in uncovering the structures mediating the highest levels of performance in domains of expertise, such as science, arts, games and professions, should consider the methods and theories offered by the expert-performance framework and be prepared to make some exciting discoveries that will contribute to our increasing understanding of the development of expertise and expert performance.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin*, *102*, 3–27. http://dx.doi.org/10.1037/0033-2909.102.1.3.

Ackerman, P. L. (2013). Intelligence and expertise. In R. J. Sternberg, & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 847–860). New York: Cambridge University Press.

Ackerman, P. L. (2014). Nonsense, common sense, and science of expert performance: Talent and individual differences. *Intelligence* (this issue).

Anastasi, A. (1958). Heredity, environment, and the question "How?". *Psychological Review*, *65*, 197–208. http://dx.doi.org/10.1037/h0044895.

Anastasi, A. (1974). Commentary on precocity project. In J. C. Stanley, D. Keating, & L. H. Fox (Eds.), *Mathematical talent: discovery, description and development* (pp. 87–100). Baltimore, MD: Johns Hopkins University Press.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369–406.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Bartkovich, K. G., & Mezynski, K. (1981). Fast-paced precalculus mathematics for talented junior high students: Two recent SMPY programs. *Gifted Child Quarterly*, *25*, 73–80. http://dx.doi.org/10.1177/001698628102500206.

Bayer, A. E., & Folger, J. (1966). Some correlates of a citation measure of productivity in science. *Sociology of Education*, *39*, 381–390.

Bédard, J., & Chi, M. T. H. (1993). Expertise in auditing. *Auditing*, *12*, 1–25 (Suppl.).

Benbow, C. P., & Arjmand, O. (1990). Predictors of high academic achievement in mathematics and science by mathematically talented students: A longitudinal study. *Journal of Educational Psychology*, *82*, 430–441. http://dx.doi.org/10.1037/0022-0663.82.3.430.

Benbow, C. P., Lubinski, D., & Suchy, B. (1996). The impact of SMPY's educational programs from the perspective of the participant. In C. P. Benbow, & D. Lubinski (Eds.), *Intellectual talent: Psychometric and social issues* (pp. 266–300). Baltimore: Johns Hopkins University Press.

Berri, D. J., & Simmons, R. (2011). Catching a draft: On the process of selecting quarterbacks in the National Football League amateur draft. *Journal of Productivity Analysis*, *35*, 37–49. http://dx.doi.org/10.1007/s11123-009-0154-6.

Bilalić, M., McLeod, P., & Gobet, F. (2007). Does chess need intelligence? A study with young chess players. *Intelligence*, *35*(5), 457–470.

Bloom, B. S. (1985). Generalizations about talent development. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 507–549). New York: Ballantine Books.

Bogl, L. H., Pietilainen, K. H., Rissanen, A., & Kaprio, J. (2009). Improving the accuracy of selfreports on diet and physical exercise: the co-twin control method. *Twin Research and Human Genetics*, *12*, 531–540.

Bouchard, T. J., Jr., & Lykken, D. T. (1999). Genetic and environmental influence on correlates of creativity. In N. Colangelo, & S. G. Assouline (Eds.), *Talent Development III: Proceedings from The 1995 Henry B. & Jocelyn Wallace National Symposium on Talent Development* (pp. 81–97). Scottsdale, AZ: Gifted Psychology Press.

Brody, L. E., & Benbow, C. P. (1987). Accelerative strategies: How effective are they for the gifted? *Gifted Child Quarterly*, *31*, 105–110.

Bühler, W. K. (1981). *Gauss: A biographical study*. New York: Springer.

Butcher, L. M., Davis, O. S., Craig, I. W., & Plomin, R. (2008). Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500 K single nucleotide polymorphism microarrays. *Genes, Brain, and Behavior*, *7*, 435–446. http://dx.doi.org/10.1111/j.1601-183X.2007.00368.x.

Campbell, J. R. (1996). Early identification of mathematics talent has long-term positive consequences for career contributions. *International Journal of Educational Research*, *25*, 497–522.

Campitelli, G., & Gobet, F. (2008). The role of practice in chess: A longitudinal study. *Learning and Individual Differences*, *18*, 446–458.

Chabris, C. F., Hebert, B. M., Benjamin, D. J., Beauchamp, J., Cesarini, D., van der Loos, M., & Laibson, D. (2012). Most reported genetic associations with general intelligence are probably false positives. *Psychological Science*, *23*, 1314–1323. http://dx.doi.org/10.1177/0956797611435528.

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.

Choudhry, N. K., Fletcher, R. H., & Soumerai, S. B. (2005). The relationship between clinical experience and health care. *Annals of Internal Medicine*, *142*, 260–273. http://dx.doi.org/10.7326/0003-4819-142-4-200502150-00008.

Clark, E. V. (2009). *First language acquisition* (2nd ed.)New York: Cambridge University Press.

Cleary, C. (August 30th). No shortage of fairy-tale finishes at worlds. *New York Times*, 137.

Cole, J. R., & Cole, S. (1973). *Social stratification in science*. Chicago, IL: Univeristy of Chicago Press.

Copp, E. F. (1916). Musical ability. *Journal of Heredity*, *7*, 297–305.

Côté, J., Ericsson, K. A., & Law, M. (2005). Tracing the development of athletes using retrospective interview methods: A proposed interview and validation procedure for reported information. *Journal of Applied Sport Psychology*, *17*, 1–19.

Coyle, T. R., Purcell, J. M., Snyder, A. C., & Kochunov, P. (2013). Non-g residuals of the SAT and ACT predict specific abilities. *Intelligence*, *41*, 114–120. http://dx.doi.org/10.1016/j.intell.2012.12.001.

Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S. E., Liewald, D., & Deary, I. J. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry*, *16*, 996–1005. http://dx.doi.org/10.1038/mp.2011.85.

de Bruin, A. B. H., Kok, E. M., Leppink, J., & Camp, G. (2014). Practice, intelligence, and enjoyment in novice chess players: A prospective study at the earliest stage of a chess career. *Intelligence* (this issue).

Der Spiegel. (1987). Genieblitz und Blackouts: Der Spiegel testete Intelligenz, Gedachtnis und Schachkunst Garri Kasparows. *Der Spiegel*, December 21 (pp. 126–140).

Doll, J., & Mayr, U. (1987). Intelligenz und Schachleistung—eine Untersuchung an Schachexperten {Intelligence and performance in chess—A study of chess experts}. *Psychologische Beiträge*, *29*, 270–289.

Dollinger, S. J., & Leong, F. T. L. (1993). Volunteer bias and the five-factor model. *The Journal of Psychology*, *127*, 29–36.

Eiduson, B. T. (1962). *Scientists; Their psychological world.* New York: Basic books.

Ericsson, K. A. (1987). Theoretical implications from protocol analysis on testing and measurement. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology in testing* (pp. 191–226). Hillsdale, N. J.: Erlbaum.

Ericsson, K. A. (1988). Analysis of memory performance in terms of memory skill. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*, *Vol. 4*. (pp. 137–179)Hillsdale, NJ: Erlbaum.

Ericsson, K. A. (1990). Peak performance and age: An examination of peak performance in sports. In P. B. Baltes, & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 164–195). New York: Cambridge University Press.

Ericsson, K. A. (2000). How experts attain and maintain superior performance: Implications for the enhancement of skilled performance in older individuals. *Journal of Aging and Physical Activity*, *8*, 346–352.

Ericsson, K. A. (2002). Attaining excellence through deliberate practice: Insights from the study of expert performance. In M. Ferrari (Ed.), *The pursuit of excellence in education* (pp. 21–55). Hillsdale, N. J.: Erlbaum.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*, S70–S81.

Ericsson, K. A. (2006a). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 685–706). Cambridge, UK: Cambridge University Press.

Ericsson, K. A. (2006b). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative task. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 223–242). Cambridge, UK: Cambridge University Press.

Ericsson, K. A. (2007a). 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. (2007b). An expert-performance perspective on medical expertise: Study superior clinical performance rather than experienced clinicians! *Medical Education*, *41*, 1124–1130.

Ericsson, K. A. (2007c). Deliberate practice and the modifiability of body and mind: A reply to the commentaries. *International Journal of Sport Psychology*, *38*, 109–123.

Ericsson, K. A. (2009). Enhancing the development of professional performance: Implications from the study of deliberate practice. In K. A. Ericsson (Ed.), *The development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 405–431). New York: Cambridge University Press.

Ericsson, K. A. (2013a). 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. (2013b). Creative genius: A view from the expert-performance approach. In D. K. Simonton (Ed.), *The Wiley-Blackwell Handbook of Genius* (in press).

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725–747.

Ericsson, K. A., Charness, N., Feltovich, P., & Hoffman, R. R. (Eds.). (2006). *Cambridge handbook of expertise and expert performance*. Cambridge, UK: Cambridge University Press.

Ericsson, K. A., & Faivre, I. (1988). What's exceptional about exceptional abilities? In L. K. Obler, & D. Fein (Eds.), *The exceptional brain: Neuropsychology of talent and special abilities* (pp. 436–473). New York: Guilford.

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

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptations to task constraints. *Annual Review of Psychology*, *47*, 273–305.

Ericsson, K. A., & Moxley, J. H. (2012). A critique of Howard's argument for innate limits of 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 of the Education of the Gifted*, *28*, 287–311.

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

Ericsson, K. A., Whyte, J., & Ward, P. (2007). Expert performance in nursing: Reviewing research on expertise in nursing within the framework of the expert-performance approach. *Advances in Nursing Science*, *30*, E58–E71.

Eynon, N., Alves, A. J., Yamin, C., Sagiv, M., Duarte, J. A., Oliveira, J., et al. (2009). Is there an ACE ID—ACTN3 R577X polymorphisms interaction that influences sprint performance? *International Journal of Sports Medicine*, *30*, 888–891.

Fehrenbach, C. R. (1991). Gifted/average readers: Do they use the same reading strategies? *Gifted Child Quarterly*, *35*, 125–127.

Feist, G. J. (2006). *The psychology of science and the origins of the scientific mind.* New Haven, CT: Yale University Press.

Feist, G. J., & Barron, F. X. (2003). Predicting creativity from early to late adulthood: Intellect, potential, and personality. *Journal of Research in Personality*, *37*, 62–88.

Fitts, P., & Posner, M. I. (1967). *Human performance.* Monterey, CA: Brooks/Cole.

Fox, L. H. (1974). A mathematics program for fostering precocious achievement. In J. C. Stanley, D. Keating, & L. H. Fox (Eds.), *Mathematical talent: discovery, description and development* (pp. 101–125). Baltimore, MD: Johns Hopkins University Press.

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. Sir (1979). *Hereditary genius: An inquiry into its laws and consequences.* London: Julian Friedman Publishers (Originally published in 1869).

Gardner, H. (1994). *The arts and human development: A psychological study of the artistic process.* New York: Basic books (Originally published in 1973).

Gardner, H. (1995). Why would anyone become an expert? *American Psychologist*, *50*, 802–803. http://dx.doi.org/10.1037/0003-066X.50.9.802.

Gibson, J., & Light, P. (1967). Intelligence among university scientists. *Nature*, *213*, 441–443. http://dx.doi.org/10.1038/213441a0.

Glover, H. (2013). *Helen Glover biography*. Great Britain Rowing Team (Accessed on November 8th 2013. http://www.britishrowing.org/gb-rowing-team/biographies/helen-glover)

Gobbo, C., & Chi, M. T. H. (1986). How knowledge is structured and used by expert and novice children. *Cognitive Development*, *1*, 221–237.

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.

Gottfried, A. W., Gottfried, A. E., Bathurst, K., & Guerin, D. W. (1994). *Gifted IQ: Early developmental aspects.* New York: Plenum.

Grabner, R. H. (2014). The role of intelligence for performance in the prototypical expertise domain of chess. *Intelligence* (this issue).

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.

Greene, J. A., Moos, D. C., Azevedo, R., & Winters, F. I. (2008). Exploring differences between gifted and grade-level students' use of self-regulatory learning processes with hypermedia. *Computers & Education*, *50*, 1069–1083. http://dx.doi.org/10.1016/j.compedu.2006.10.004.

Gromko, J. E. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education*, *52*, 6–15.

Hagen, S. L., Cremaschi, A., & Himonides, C. S. (2012). Effects of extended practice with computerized eye guides for sight-reading in collegiate-level class piano. *Journal of Music, Technology & Education*, *5*, 229–239.

Hallam, S. (1998). The predictors of achievement and dropout in instrumental tuition. *Psychology of Music*, *26*, 116–132. http://dx.doi.org/10.1177/0305735698262002.

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* (this issue).

Hamilton, W. (1986). Physical prerequisites for ballet. *Journal of Musculoskeletal Medicine*, *3*, 61–66.

Hamilton, W. G. (1988). Foot and ankle injuries in dancers. *Clinics in Sports Medicine*, *7*, 143–173.

Hanushek, E. A., & Rivkin, S. J. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, *100*, 267–271. http://dx.doi.org/10.1257/aer.100.2.267.

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001

Harmon, L. R. (1961). High school backgrounds of science doctorates. *Science*, *133*, 679–688. http://dx.doi.org/10.1126/science.133.3454.679.

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.

Hegarty, M., Keehner, M., Khooshabeh, P., & Montello, D. R. (2009). How spatial abilities enhance and are enhanced by, dental education. *Learning and Individual Differences*, *19*, 61–70. http://dx.doi.org/10.1016/j.lindif.2008.04.006.

Helmbold, N., Rammsayer, T., & Altenmuller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. *Journal of Individual Differences*, *26*, 74–85. http://dx.doi.org/10.1027/1614-0001.26.2.74.

Holmes, F. L. (1996). Scientific creativity in biomedical research: Some case histories. *European Journal of Clinical Chemistry and Clinical Biochemistry*, *34*, 941–948. http://dx.doi.org/10.1515/cclm.1996.34.12.941.

Horn, J., & Masunaga, H. (2006). A merging theory of expertise and intelligence. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 587–611). Cambridge, UK: Cambridge University Press.

Horn, J. L., & McArdle, J. J. (2007). Understanding human intelligence in factor analysis since Spearman. In R. Cudek, & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* (pp. 205–247). Mahwah, N.J.: Lawrence Erlbaum Associates.

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

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.

Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences*, *21*, 399–442.

Hu, Y., & Ericsson, K. A. (2012). Memorization and recall of very long lists accounted for within the Long-Term Working Memory framework. *Cognitive Psychology*, *64*, 236–266. http://dx.doi.org/10.1016/j.cogpsych.2012.01.001.

Hufstader, R. A. (1974). Predicting success in beginning instrumental music through use of selected tests. *Journal of Research in Music Education*, *22*, 52–57. http://dx.doi.org/10.2307/3344618.

Hutchinson, C. U., Sachs-Ericsson, N. J., & Ericsson, K. A. (2013). Generalizable aspects of the development of expertise in ballet across countries and cultures: A perspective from the expert-performance approach. *High Ability Studies*, *24*, 21–47. http://dx.doi.org/10.1080/13598139.2013.780966.

Jääskeläinen, R. (2010). Are all professionals experts? Definitions of expertise and reinterpretation of research evidence in process studies. In G. M. Shreve, & E. Angelone (Eds.), *Translation and cognition* (pp. 213–227). Amsterdam, the Netherlands: John Benjamins Publishing Company.

Jensen, A. R. (1987). The g beyond factor analysis. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology in testing* (pp. 87–142). Hillsdale, N. J.: Erlbaum.

Jensen, A. R. (1990). Speed of information processing in a calculating prodigy. *Intelligence*, *14*, 259–274. http://dx.doi.org/10.1016/0160-2896(90)90019-P.

Johnson, B. J., Tenenbaum, G., & Edmonds, W. A. (2006). Adaptation to physical and emotional demanding conditions: The role of deliberate practice. *High Abilities Studies*, *I7*, 117–136.

Jung, W. H., Kim, S. N., Lee, T. Y., Jang, J. H., Choi, C. -H., Kang, D. -H., & Kwon, J. S. (2013). Exploring the brains of *Baduk* (*Go*) experts: Gray matter morphometry, resting-state functional connectivity, and graph theoretical analysis. *Frontiers in Human Neuroscience*, *7*(633), 1–16.

Keehner, M., Lippa, Y., Montello, D. R., Tendick, F., & Hegarty, M. (2006). Learning a spatial skill for surgery: How the contributions of abilities change with practice. *Applied Cognitive Psychology*, *20*, 487–503.

Klinedinst, R. E. (1991). Predicting performance achievement and retention of fifth-grade instrumental students. *Journal of Research in Music Education*, *39*, 225–238. http://dx.doi.org/10.2307/3344722.

Kopiez, R., & Lee, J. L. (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. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, *125*, 331–359.

Kyvik, K. O., Green, A., & Beck-Nielsen, H. (1995). The new Danish twin register: Establishment and analysis of twinning rates. *International Journal of Epidemiology*, *24*, 589–596.

Law, K. S., Wong, C. S., Huang, G. H., & Li, X. (2008). The effects of emotional intelligence on job performance and life satisfaction for the research and development scientists in China. *Asia Pacific Journal of Management*, *25*, 51–69. http://dx.doi.org/10.1007/s10490-007-9062-3.

Lee, B., Park, J. Y., Jung, W. H., Kim, H. S., Oh, J. S., & Kwon, J. S. (2010). White matter neuroplastic changes in long-term trained players of the game of "Baduk" (GO): A voxel-based diffusion-tensor imaging study. *NeuroImage*, *52*, 9–19. http://dx.doi.org/10.1016/j.neuroimage.2010.04.014.

Lehmann, A. C., & Ericsson, K. A. (1998). The historical development of domains of expertise: Performance standards and innovations in music. In A. Steptoe (Ed.), *Genius and the mind* (pp. 67–94). Oxford, UK: Oxford University Press.

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology*, *12*(2), 182–195. http://dx.doi.org/10.1037/h0094108.

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

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.

Lykken, D. T. (1982). Research with twins: The concept of emergenesis. *Psychophysiology*, *19*, 361–373. http://dx.doi.org/10.1111/j.1469-8986.1982.tb02489.x.

Lykken, D. T. (1998). The genetics of genius. In A. Steptoe (Ed.), *Genius and the mind* (pp. 15–37). Oxford, UK: Oxford University Press.

Lynn, R., & Gault, A. (1986). The relation of musical ability to general intelligence and the major primaries. *Research in Education*, *36*, 59–64.

Lynn, R., Wilson, G., & Gault, A. (1989). Simple musical tests as measures of Spearman's g. *Personality and Individual Differences*, *10*, 25–28.

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.

Ma, F., Yang, Y., Li, X., Zhou, F., Gao, C., Li, M., & Gao, L. (2013). The association of sport performance with ACE and ACTN3 genetic polymorphisms: A systematic review and meta-analysis. *PLoS ONE*, *8*, e54685. http://dx.doi.org/10.1371/journal.pone.0054685.

MacArthur, D. G., & North, K. N. (2005). Genes and human elite athletic performance. *Human Genetics*, *116*, 331–339. http://dx.doi.org/10.1007/s00439-005-1261-8.

MacKinnon, D. W. (1962). The nature and nurture of creative talent. *American Psychologist*, *17*, 484–495.

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

McCaffery, J. M., Papandonatos, G. D., Bond, D. S., Lyons, M. J., & Wing, R. R. (2009). Gene X environment interaction of vigorous exercise and body mass index among male Vietnam-era twins. *American Journal of Clinical Nutrition*, *89*, 1011–1018. http://dx.doi.org/10.3945/ajcn.2008.27170.

McKelvie, S. J. (1989). The Wonderlic Personnel Test: Reliability and validity in an academic setting. *Psychological Reports*, *65*, 161–162.

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.

Mishra, J. (2013). Improving sightreading accuracy: A meta-analysis. *Psychology in Music*. http://dx.doi.org/10.1177/0305735612463770 (in press).

Mitchell, F. D. (1907). Mathematical prodigies. *American Journal of Psychology*, *18*, 61–143. http://dx.doi.org/10.2307/1412172.

Montgomery, E. C., Kunik, M. F., Wilson, N., Stanley, M. A., & Weiss, B. (2010). Can paraprofessionals deliver cognitive-behavioral therapy to treat anxiety and depressive symptoms? *Bulletin of the Menninger Clinic*, *74*, 45–62. http://dx.doi.org/10.1521/bumc.2010.74.1.45.

Nokelainen, P., Tirri, K., & Campbell, J. R. (2004). Cross-cultural predictors of mathematical talent and academic productivity. *High Ability Studies*, *15*, 229–242. http://dx.doi.org/10.1080/1359813042000314790.

Nokelainen, P., Tirri, K., Campbell, J. R., & Walberg, H. (2007). Factors that contribute to or hinder academic productivity: Comparing two groups of most and least successful Olympians. *Educational Research and Evaluation*, *13*, 483–500. http://dx.doi.org/10.1080/13803610701785931.

Pelliccia, A., Maron, B. J., de Luca, R., di Paolo, F. M., Spataro, A., & Culasso, F. (2002). Remodeling of left ventricular hypertrophy in elite athletes after long-term deconditioning. *Circulation*, *105*, 944–949.

Pitsiladis, Y., Wang, G., Wolfarth, B., Scott, R., Fuku, N., Mikami, E., et al. (2013). Genomics of elite sporting performance: What little we know and necessary advances. *British Journal of Sports Medicine*, *47*, 550–555. http://dx.doi.org/10.1136/bjsports-2013-092400.

Plomin, R. (2012). Genetics: How intelligence changes with age. *Nature*, *482*, 165–166. http://dx.doi.org/10.1038/482165a.

Plomin, R., Owen, M. J., & McGuffin, P. (1994). The genetic basis of complex human behaviors. *Science*, *264*, 1733–1739. http://dx.doi.org/10.1126/science.8209254.

Plomin, R., Shakeshaft, N. G., McMillan, A., & Trzaskowski, M. (2014). Nature, nurture, and expertise. *Intelligence* (this issue).

Renzulli, J. S., & McGreevy, A. M. (1986). Twins included and not included in special programs for the gifted. *Roeper Review*, *9*, 120–127.

Roe, A. (1953a). *The making of a scientist.* New York: Dodd, Mead.

Roe, A. (1953b). A psychological study of eminent psychologists and anthropologists, and a comparison with biological and physical scientists. *Psychological Monographs*, *67* (Whole No. 352).

Root-Bernstein, R., Allen, L., Beach, L., Bhadula, R., Fast, J., & Weinlanders, S. (2008). Arts foster scientific success: Avocations of Nobel, National Academy, Royal Society, and Sigma Xi members. *Journal of the Psychology of Science and Technology*, *1*, 51–63.

Root-Bernstein, R. S., Bernstein, M., & Garnier, H. (1993). Identification of scientists making long-term, high impact contributions, with notes on their methods of working. *Creativity Research Journal*, *6*, 329–343.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103*, 734–760.

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., Ruthsatz, K., & Stephens, K. R. (2014). Putting practice into perspective: Child prodigies as evidence of innate talent. *Intelligence* (this issue).

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. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162–173.

Scott, R. A., Irving, R., Irwin, L., Morrison, E., Charlton, V., Austin, K., et al. (2010). ACTN3 and ACE genotypes in elite Jamaican and US sprinters. *Medicine and Science in Sports and Exercise*, *42*, 107–112.

Seshadri, S., DeStefano, A. L., Au, R., Massaro, J. M., Beiser, A. S., Kelly-Hayes, M., & Wolf, P. A. (2007). Genetic correlates of brain aging on MRI and cognitive test measures: A genome-wide association and linkage analysis in the Framingham Study. *BMC Medical Genetics*, *8*, S15.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.

Shuter, R. (1968). *The psychology of musical ability.* London: Methuen and Co.LTD.

Simonton, D. K. (1991). Career landmarks in science: Individual differences and interdisciplinary contrasts. *Developmental Psychology*, *27*, 119–130.

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. (2005). Genetics of giftedness: The implications of an emergenic-epigenetic model. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 312–326) (2nd ed.). Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511610455.018.

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.

Simonton, D. K. (2014). Creative performance, expertise acquisition, individual differences, and developmental antecedents: An integrative research agenda. *Intelligence* (this issue).

Spengler, P., White, M., Ægisd´ ottir, S., Maugherman, A., Anderson, L., Cook, R., & Rush, J. (2009). The meta-analysis of clinical judgment project: Effects of experience on judgment accuracy. *The Counseling Psychologist*, *37*, 350–398.

Stanley, J. C. (1977). Rationale for the Study of Mathematically Precocious Youth (SMPY) during its first five years of promoting educational acceleration. In J. C. Stanley, W. C. George, & C. H. Solano (Eds.), *The gifted and the creative: A fifty-year perspective* (pp. 75–112). Baltimore, MD, US: Johns Hopkins University Press.

Swiatek, M. A., & Benbow, C. P. (1991). A ten-year longitudinal follow-up of participants in a fast-paced mathematics course. *Journal for Research in Mathematics Education*, *22*, 138–150. http://dx.doi.org/10.2307/749590.

Tarchi, C. (2010). Reading comprehension of informative texts in secondary school: A focus on direct and indirect effects of reader's prior knowledge. *Learning and Individual Differences*, *20*, 415–420.

Tetlock, P. E. (2005). *Expert political judgment: How good is it? How can we know?* Princeton, NJ: Princeton Univ. Press.

Threlfall, J., & Hargreaves, M. (2008). The problem-solving methods of mathematically gifted and older average-attaining students. *High Ability Studies*, *19*(1), 83–98. http://dx.doi.org/10.1080/13598130801990967.

Treffert, D. A. (1989). *Extraordinary people: Understanding "idiots savants.".* New York: Harper & Row.

Tucker, R., & Collins, M. (2012). What makes champions? A review of the relative contribution of genes and training to sporting success. *British Journal of Sports Medicine*, *46*, 555–561. http://dx.doi.org/10.1136/bjsports-2011-090548.

Unterrainer, J. M., Kaller, C. P., Halsband, U., & Rahm, B. (2006). Planning ability and chess. A comparison of chess and non-chess players on the Tower of London task. *British Journal of Psychology*, *97*, 299–311.

Unterrainer, J. M., Kaller, C. P., Leonhart, R., & Rahm, B. (2011). Revising superior planning performance in chess players: The impact of time restriction and motivation aspects. *American Journal of Psychology*, *124*, 213–225. http://dx.doi.org/10.5406/amerjpsyc.124.2.0213.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*, 352–402.

VanTassel-Baska, J. (1998). A critique of the talent searches. *Journal for Secondary Gifted Education*, *9*, 139–144.

Vinchur, A. J., Schippmann, J. S., Switzer, F. S., III, & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology*, *83*, 586–597.

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.

Wai, J. (2013). Investigating America's elite: Cognitive ability, education, and sex differences. *Intelligence*, *41*, 203–211.

Wai, J. (2014). Experts are born, then made: Combining prospective and retrospective longitudinal data shows that cognitive ability matters. *Intelligence* (this issue).

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and itsrelation to STEM educational dose: A 25-year longitudinal study. *Journalof Educational Psychology*, *102*, 860–871.

Walker, D. E. (1955). The relationship between creativity and selected test behaviors for chemists and mathematicians. Unpublished doctoral dissertation, University of Chicago.

Wanzel, K. R., Hamstra, S. J., Caminiti, M. F., Anastakis, D. J., Grober, E. D., & Reznick, R. K. (2003). Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. *Surgery*, *134*, 750–757.

Waters, A. J., Gobet, F., & Leyden, G. (2002). Visuospatial abilities of chess players. *British Journal of Psychology*, *93*, 557–565.

Wellington, C. (2013). *Chrissie Wellington's running secrets.* United Kingdom: National Health Services (Accessed November 8th, 2013 http://www.nhs.uk/Livewell/c25k/Pages/chrissie-wellington.aspx)

Wilber, R. L., & Pitsiladis, Y. P. (2012). Kenyan and Ethiopian distance runners: what makes them so good? *International Journal of Sports Physiology & Performance*, *7*, 92–102.

Williams, P. T., Blanche, P. J., & Krauss, R. M. (2005). Behavioral versus genetic correlates of lipoproteins and adiposity in identical twins discordant for exercise. *Circulation*, *112*, 350–356.

Wong, A. C. -N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert but which type? *Psychological Science*, *20*. http://dx.doi.org/10.1177/0956797611401753 (1108logica).

Woollett, K., & Maguire, E. A. (2011). Acquiring "the Knowledge" of London's layout drives structural brain changes. *Current Biology*, *21*, 2109–2114.

Wu, W. T. (1996). Growing up in Taiwan: The impact of environment influence on the math Olympians. In J. R. Campbell (Ed.), *Cross-national retrospective studies of mathematics Olympians (pp. 525–534). International Journal of Educational Research*, *25*. (pp. 523–534).

Young, W. T. (1971). The role of musical aptitude, intelligence, and academic achievement in predicting the musical attainment of elementary instrumental music students. *Journal of Research in Music Education*, *19*, 385–398.

Young, B. W., & Salmela, J. H. (2010). Examination of practice activities related to the acquisition of elite performance in Canadian middle distance running. *International Journal of Sport Psychology*, *41*, 73–90.

Please cite this article as: Ericsson, K.A., Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.12.001