Psychological Bulletin

Rethinking Expertise: A Multifactorial Gene-Environment Interaction Model of Expert Performance
Fredrik Ullén, David Zachary Hambrick, and Miriam Anna Mosing

CITATION
Rethinking Expertise: A Multifactorial Gene–Environment Interaction Model of Expert Performance

Fredrik Ullén
Karolinska Institutet

Miriam Anna Mosing
Karolinska Institutet

David Zachary Hambrick
Michigan State University

Scientific interest in expertise—superior performance within a specific domain—has a long history in psychology. Although there is a broad consensus that a long period of practice is essential for expertise, a long-standing controversy in the field concerns the importance of other variables such as cognitive abilities and genetic factors. According to the influential deliberate practice theory, expert performance is essentially limited by a single variable: the amount of deliberate practice an individual has accumulated. Here, we provide a review of the literature on deliberate practice, expert performance, and its neural correlates. A particular emphasis is on recent studies indicating that expertise is related to numerous traits other than practice as well as genetic factors. We argue that deliberate practice theory is unable to account for major recent findings relating to expertise and expert performance, and propose an alternative multifactorial gene–environment interaction model of expertise, which provides an adequate explanation for the available empirical data and may serve as a useful framework for future empirical and theoretical work on expert performance.

Keywords: expertise, learning, plasticity, intelligence, genetics

Supplemental materials: http://dx.doi.org/10.1037/bul0000033.supp

Learning and neural plasticity are ubiquitous in the animal kingdom, but the learning capacity of humans appears remarkable compared with most other animals. One reason for this is that humans can engage in both practice—persistent, goal-directed activities designed to increase knowledge and skills—and teaching, the active training of learners by more experienced individuals (Premack, 2007). Together, practice and teaching enable individuals to gradually acquire highly specialized competencies, which are needed for achieving consistently superior levels of performance within a particular domain, that is, to become experts. The variety of different forms of vocational and avocational expertise across cultures is further testament to the flexibility of human learning. On domain-specific tasks, the performance differences between experts and nonexperts can be massive (Ericsson & Lehmann, 1996), from both a statistical and a practical perspective. A striking feature of human societies is thus not just a complex division of labor but also a vast range of interindividual variability in skills and abilities.

Scientific interest in expertise and related topics has a long history. Indeed, psychologists have been interested in the origins of expertise for as long as psychology has been a field. A long-standing controversy in the area concerns the relative importance of nature and nurture for achievement and expertise. In the late 1800s, Francis Galton (1869) analyzed family trees of eminent scientists, musicians, writers, and other professionals, and noted that accomplishment tends to run in families. Based on these observations, Galton concluded that “natural ability” is important for eminence and achievement. More than half a century later, the behaviorist John Watson famously articulated the alternative environmental view when he speculated that, in a hypothetical scenario in which the environment is completely controlled, he could take any infant at random and train him to become “any type of specialist . . . regardless of his talents” (Watson, 1998/1924, p. 82).

More recently, scientists from a number of different research traditions have used a wide variety of methods and approaches to study the origins and underpinnings of expertise, ranging from historiometric and biographical studies of high-achieving individuals, to sociological, psychological, computational, and neurobiological studies of exceptional performance and its mechanisms (Ericsson, Charness, Feltovich, & Hoffman, 2006; Ericsson & Smith, 1991; Eysenck, 1995; Galton, 1869; Gardner, 1993; Gobet, DeVoogt, & Retschitzki, 2004; Howe, Davidson, & Sloboda, 1998; Simonton, 2004; Sternberg & Grigorenko, 2003; Terman, 1925).

A paradigm that has been particularly influential in psychological research on expertise over the past two decades is the expert
The expert performance framework emphasizes the importance of studying expert performance using objective measures of performance on standardized, representative tasks rather than more indirect indices of expertise, such as estimates by raters, credentials, or educational level (Ericsson & Lehmann, 1996; Ericsson & Ward, 2007). The roots of this research tradition can be traced back to early studies of complex skill learning (Bryan & Harter, 1897, 1899; for historical reviews, see, e.g., Adams, 1987, and Feltonvich et al., 2006), but a decisive influence came from the computational approach to studies of information processing in human problem solving proposed by Allen Newell, Herbert Simon, and others (Newell & Simon, 1972). In seminal studies, based on earlier work by de Groot (1965/1978), this approach was applied to the analysis of cognitive processes underlying chess playing in expert chess players by Simon and Chase (1973; Chase & Simon, 1973b). From the beginning, there was thus a strong interest in understanding how information processing in experts differs from that of novices.

Although fundamentally different from classical behavioralism in its focus on the internal representations and processes underlying skilled behavior, in one important sense, the expert performance framework can be seen as a continuation of the tradition from Watson: its strong emphasis on the importance of acquired characteristics for expertise (Ericsson, Krampe, & Tesch-Römer, 1993). In particular, this framework proposes that expert performance reflects the accumulated amount of deliberate practice, defined as explicit, effortful, goal-directed activities that are specifically designed to improve performance. The expert performance framework further proposes that merely spending time with a summary of some general observations in the field that are level performance and achievement. The present review begins with a summary of some general observations in the field that are well supported by data and that are of particular theoretical interest because they appear to hold true across different domains of expertise. We then review findings that seriously challenge deliberate practice theory, focusing in particular on findings from recent studies demonstrating that variables other than deliberate practice account for individual differences in expertise. In the final section, we argue that deliberate practice theory fails to account for major recent findings relating to expertise and expert performance, and propose an alternative framework for understanding and investigating expertise—the multifactorial gene–environment interaction model (MGIM).

The literature considered in the present review includes studies on expertise and deliberate practice identified from earlier reviews and books, as well as from systematic literature searches in Web of Science using the search terms “expertise” OR “expert performance,” “imaging,” and “transfer,” and the Web of Sciences categories “neurosciences,” “neuroimaging,” “psychology,” “behavioral sciences,” and “genetics heredity.” Details on the literature search are provided as online supplemental materials.

The breadth of the field makes it impossible to perform an exhaustive review of all original studies in a single review article. Our main focus is therefore on studies from the last few decades. Finally, it should be stressed that we primarily consider a number of expertise domains that have been extensively studied in the deliberate practice literature, including music, dance, board games, video games, and sports. Other forms of expertise, such as academic expertise, will only be discussed more tangentially.

### The Neuropsychology of Expertise

#### Practice, Expertise, and Brain Anatomy

A large amount of empirical data supports that cumulative estimates of lifetime deliberate practice are positively associated with measures of expertise (Hambrick, Oswald, et al., 2014; Macnamara, Hambrick, & Oswald, 2014; Platz, Kopiez, Lehmann, & Wolf, 2014). A meta-analysis of 88 published studies corroborates this basic finding (Macnamara et al., 2014); for a comprehensive bibliography, we refer the reader to the supplementary material of that article. Correlations between practice and expertise have been demonstrated using various measures of both practice (interview data, questionnaires, logs) and expertise (objective measures of expert performance, ratings, professional status). The literature covers a wide range of domains, including arts, sports, games, education, and professions (Ericsson et al., 1993; Hambrick, Oswald, et al., 2014; C. U. Hutchinson, Sachs-Ericsson, & Ericsson, 2013; Macnamara et al., 2014; Platz et al., 2014; Tucker & Collins, 2012), Macnamara and colleagues (2014) found an overall meta-analytic average correlation between deliberate practice and expert performance of $r = .35$ across domains. However, the effect size of the relation varied between domains. The percentage of variance in expert performance explained by deliberate practice was 1% for professions, 4% for education, 18% for sports, 21% for music, and 26% for games. In another meta-analysis, Platz and colleagues (2014) found that deliberate practice could account for 36% of the variance in musical expertise. For further discussion of these associations, see the Challenges to Deliberate Practice Theory section.

Common sense and everyday experience make it tempting to assume that correlations between deliberate practice and expert performance reflect causal effects of the former on the latter. At
the same time, correlational findings are often consistent with many underlying causal mechanisms, and recent findings suggest that alternative scenarios also have to be considered. Evidence has emerged that at least some components of expert performance are not causally influenced by deliberate practice (see the Challenges to Deliberate Practice Theory section). However, studies investigating the acquisition of expertise longitudinally based on single cases (Ericsson, Chase, & Faloon, 1980), or observational (Campitelli & Gobet, 2008; de Bruin, Kok, Leppink, & Camp, 2014; Howard, 2012; Hyde et al., 2009; Jabbusch, Alpers, Kopiez, Vauth, & Altenmüller, 2009; McPherson, 2005) and randomized (Bilhartz, Bruhn, & Olson, 1999; Scholz, Klein, Behrens, & Johansen-Berg, 2009) group designs suggest that, as expected, long-term deliberate practice is an essential causal factor behind many aspects of expert performance.

Many studies have also investigated expertise using neuroimaging techniques. This literature, which mainly consists of correlational studies, further adds to the evidence that deliberate practice is an important predictor of expertise, and suggests that individual differences in regional brain anatomy may in part mediate associations found between deliberate practice and expert performance on the behavioral level. One of the most frequently studied domains in this literature is musicians (Münthe, Altenmüller, & Jäncke, 2002; Zatorre, Fields, & Johansen-Berg, 2012). Correlations between musical expertise and regional brain anatomy have been reported for motor and premotor areas (Amunts et al., 1997; Bangert & Schlaug, 2006; Bermudez, Lerch, Evans, & Zatorre, 2009; Gaser & Schlaug, 2003; Han et al., 2009; Hyde et al., 2009), Broca’s area (Abdul-Kareem, Stancak, Parkes, & Sluming, 2011; Gaser & Schlaug, 2003; Sluming et al., 2002), auditory areas (Bermudez et al., 2009; Elmer, Hänggi, Meyer, & Jancke, 2013; Fauvel et al., 2014; Gaser & Schlaug, 2003; Hyde et al., 2009; Pantev & Herholz, 2011; Schneider et al., 2002; Seifert-Presi1er, Parncutt, & Schneider, 2014), the cerebellum (S. Hutchinson, Lee, Gaab, & Schlaug, 2003), and white matter pathways implicated in musical performance, such as the corpus callosum (Bengtsson et al., 2005; Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995; Steele, Bailey, Zatorre, & Penhune, 2013; Vollmann et al., 2014), the superior longitudinal fasciculus (Oechslin, Van De Ville, Lazeyras, Hauert, & James, 2013), and the pyramidal tracts (Bengtsson et al., 2005; Han et al., 2009).

Several studies have also reported significant correlations between measures of practice and regional brain anatomy in brain regions that show group differences in anatomy between musicians and nonmusicians. Amunts and colleagues (1997) found a Spearman correlation of $\rho = -.63$ to -.60 between the size of motor cortex and age of commencement of musical training. Sluming and colleagues (2002) found a correlation of $r = .39$ between years of musical training and gray matter volume in Broca’s area in younger adult musicians, and Abdul-Kareem et al. (2011) reported a correlation of $r = .72$ between the gray matter volume of this region and years of performance in orchestra musicians. As another example, S. Hutchinson et al. (2003) found a correlation of $r = .60$ between cerebellar volume and lifelong intensity of training in male musicians.

Correlational data also support that brain adaptations in musicians are functionally important. Hyde and colleagues (2009) used a longitudinal design and studied associations between performance improvements on relevant behavioral tasks and changes in regional anatomy after 15 months of musical training. Performance on a manual motor sequence task correlated with anatomical changes in primary motor cortex ($r = .45$), as well as the corpus callosum ($r = .45$). Performance on an auditory musical discrimination task correlated with anatomical changes in primary auditory cortex ($r = .40$). Schneider et al. (2002) found a correlation of $r = .70$ between the gray matter volume of auditory cortex and auditory musical discrimination in a mixed sample of amateur and professional musicians.

Similar associations between expertise and regional neuroanatomy have also been found for other groups, including taxi drivers (Maguire et al., 2000; Woollett & Maguire, 2011), jugglers (Draganski et al., 2004; Gerber et al., 2014; Scholz et al., 2009), dancers (Hänggi, Koeneke, Bezzola, & Jäncke, 2010), professional simultaneous interpreters (Elmer, Hänggi, Meyer, & Jancke, 2011), painters (Lorains, Ball, & MacMahon, 2013), chess players (Hänggi, Brütsch, Siegel, & Jancke, 2014), professional car drivers (Bernardi et al., 2014), karate experts (Roberts, Bain, Day, & Husain, 2013), perfumers (Delon-Martin, Plailly, Foulupt, Veyrac, & Royet, 2013), and athletes (Jäncke, Koeneke, Hoppe, Rominger, & Hänggi, 2009; Wei, Zhang, Jiang, & Luo, 2011). Furthermore, longitudinal studies provide support for a causal effect of practice (Draganski et al., 2004; Hyde et al., 2009; Scholz et al., 2009; Woollett & Maguire, 2011). Overall, then, there is substantial evidence that deliberate practice is associated with higher levels of expertise and that this association at least partly reflects causal influences of deliberate practice on expert performance and its underlying neural substrates.

**Expert–Novice Differences in Information Processing and Functional Brain Properties**

The literature on complex skill learning is vast. However, much of this research has used relatively short training periods (hours to days). A central finding in the expertise literature is that prolonged deliberate practice, continuing for months and years, may lead to the acquisition of qualitatively different performance strategies, which enable experts to perform at levels that baffle beginners (Ericsson & Lehmman, 1996; Feltovich et al., 2006). A dramatic illustration of this is Ericsson and colleagues’ (1980) longitudinal case study of a participant (SF) who engaged in regular practice of a digit memory span task for more than 1.5 years. During this period, SF’s memory span increased from seven to almost 80 digits—an increase of more than an order of magnitude. Verbal reports and analyses of temporal response patterns showed that this remarkable improvement relied on a gradual mastery of more and more complex memorization strategies. In particular, SF used his extensive knowledge of track and running times to chunk sequences of digits and represented number sequences hierarchically as groups and supergroups of shorter subsequences.

There is evidence for extremely efficient processing of domain-specific information in many forms of expertise. One well-documented phenomenon, illustrated by SF, is that experts can represent large, complex structures as integrated units, or “chunks” (Feltovich et al., 2006). Pioneering studies on expert performance demonstrated such chunking in chess experts, who display superior memory for chess game positions compared with novices (Charness, 1976; Chase & Simon, 1973b; de Groot, 1978; Simon & Chase, 1973). Similar phenomena have been demonstrated in

Importantly, chunk formation in experts reflects functional properties and relations between the elements in a chunk. Experts’ superior memory is typically most pronounced for ecologically valid patterns, as opposed to constellations that are unlikely to appear in real life expert performance. Chess experts display superior memory for game positions from real games, but to a much lower degree for random positions, and the grouping of pieces into chunks is based on the strategic relations between pieces (Chase & Simon, 1973b; Gobet & Simon, 1996a, 1996b; Saariluoma, 1994; Simon & Chase, 1973). In bridge (Engle & Bukstel, 1978), experts have superior memory for structured (i.e., ecological), but not for unstructured, bridge hands. Experienced programmers chunk computer code efficiently, based on its functional organization; beginners utilize common-language associations (McKeithen et al., 1981). Taxi drivers show exceptional memory for street names when these are organized according to driving routes, but not for randomly organized lists (Kalakoski & Saariluoma, 2001). Musicians display superior processing of well-formed, as opposed to random, musical structures (Halpern & Bower, 1982; Kalakoski, 2007; Meinz & Salthouse, 1998; Sloboda, 1976) and process musical materials using larger units than novices (Waters et al., 1997, 1998).

Clearly, these findings suggest that domain-specific information retained in long-term memory (LTM) is important for expertise. Mechanisms allowing for efficient interactions between LTM storage and working memory (WM), increasing the functional capacity of WM for domain-specific information, could be key to expert performance. Chase and Simon (1973a, 1973b) proposed a chunking theory to account for expert performance in chess. This theory proposes that players store patterns of pieces that occur frequently in games as chunks in LTM. Game positions can then be represented in WM as chunks rather than as single pieces. Even if experts and novices have the same general WM span, the experts will, in this way, display a higher effective WM capacity for game positions in chess. This may not just boost memory performance, but could also result in higher playing skills by making the search among possible moves and their evaluation much more selective and efficient.

Later work has resulted in revisions and developments of these ideas. Two influential current models of information processing in experts are the long-term working memory theory of Ericsson and Kintsch (1995) and the template theory of Gobet and Simon (1996c). Long-term working memory theory assumes that experts acquire the ability to rapidly encode and retrieve domain-specific information through extensive deliberate practice, and the use of retrieval structures in which sets of cues that are associated with the encoded information are used to access it during expert performance (Ericsson & Kintsch, 1995). A central concept in template theory is the template—a flexible high-level pattern that, unlike a fixed chunk, can have both fixed and variable components and can include associative links, for example, to action plans and other templates (Gobet & Simon, 1996c).

At present, there is little data on how the dramatic changes in information processing implied by these theories may be implemented on a neural level. Notably, efficient use of information in LTM is a central feature of all the models that have been discussed. One possibility would thus be that the medial temporal cortex and other brain areas implicated in LTM are active in aspects of expert performance involving online manipulation of information in WM. This idea was recently developed in a review by Guida, Gobet, Tardieu, and Nicolas (2012), who have found support for activation of LTM regions during expert performance in mental calculators, mnemonists, and chess experts.

One specific temporal region, the fusiform face area (FFA), has been the focus of several studies on expertise. The FFA is involved in the processing of faces (Kanwisher, McDermott, & Chun, 1997; Kanwisher & Yovel, 2006). In some expert groups, the FFA appears to have adapted so that it is also involved in holistic processing of domain-specific visual stimuli, such as chess game positions in chess players, cars in car experts, and radiological images in radiologists (Bilalić, Langner, Ulrich, & Grootd, 2011; Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Harley et al., 2009; McGugin, Gatenby, Gore, & Gauthier, 2012). Furthermore, Bilalić and colleagues (2011) found that, in chess experts but not in novices, FFA responses to random and ecological chess positions differed. This suggests that the FFA had adapted, not just to process chess stimuli in general but also to process relations between chess pieces that are potentially relevant for expert performance.

These studies clearly show that expert performance may involve sophisticated functional adaptations in higher visual areas processing objects in the domain of expertise. However, the extent to which the FFA is evolutionarily adapted to specifically process faces, and—more generally—whether visual object processing involves modular or distributed representations in inferotemporal cortical areas, remains controversial issues (Haxby, Connolly, & Guntupalli, 2014; Haxby et al., 2001; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999; Kanwisher, 2010; Kanwisher et al., 1997; McGugin, Newton, Gore, & Gauthier, 2014).

Perceptual processing of domain-specific stimuli in experts can also induce extensive activations of brain regions involved in motor control and the analysis of actions. In general, areas reliably activated during action observation include the motor and premotor areas, the inferior frontal cortex, the superior and inferior parietal cortices, and the cerebellum (Abreu et al., 2012; Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2006; Caspers, Zilles, Laird, & Eickhoff, 2010). Several studies have demonstrated that activity in this network may differ between experts and novices during observation of domain-specific actions (Chang, 2014; Liew, Sheng, Margetis, & Aziz-Zadeh, 2013; Zatorre, Chen, & Penhune, 2007). Balser, Lorey, Pilgramm, Naumann, et al. (2014) found stronger activation of superior parietal cortex, the supplementary motor area, and the cerebellum when experts in tennis and volleyball observed tennis and volleyball serves, respectively, than when they observed serves in the nontrained sport. The same group also found that tennis players had higher activity than novices in action-observation-related brain areas during a domain-specific motor anticipation task (Balser, Lorey, Pilgramm,
Stark, et al., 2014). Similarly, Calvo-Merino, Glaser, Grèzes, Passingham, and Haggard (2005) compared ballet and capoeira dancers and found higher activity in premotor and parietal regions when the participants’ observed movements from their trained dance than when they viewed the untrained dance. Wright, Bishop, Jackson, and Abernethy (2013) found greater activation in action observation regions during visual discrimination of soccer moves in skilled than in unskilled soccer players. Pianists showed higher activity than controls in action observation areas as well as auditory areas during observation of piano playing (Haslinger et al., 2005). Activation of motor areas in experts has been observed not just during domain-specific action observation but also during perception of more abstract stimuli that are associated with actions in the domain of expertise. For instance, Woods, Hernandez, Wagner, and Beilock (2014) found higher activity in sensorimotor and premotor areas in athletes than in novices during perception of sport-related environmental sounds. Behmer and Jantzen (2011) used EEG to compare brain activity in musicians and nonmusicians during passive observation of sheet music. A significantly stronger desynchronization of the sensorimotor mu rhythm, indicating an activation of the motor system, was found in the musicians. Lyons et al. (2010) found that hockey players, compared with novices, had higher activity in motor areas during passive listening to verbal stimuli (sentences) relating to ice hockey.

These findings indicate that functional reorganizations within distributed brain networks for the planning and control of actions may be important for forms of expertise that involve sensorimotor skills. A few studies have also used fMRI to find direct evidence for altered functional connectivity in experts. Fauvel et al. (2014) reported numerous differences in resting state functional connectivity between musicians and nonmusicians, including enhanced connections between temporal and frontal regions that could reflect adaptations in neural circuits for auditory-motor integration. Pinho, de Manzano, Fransson, Eriksson, and Ullén (2014) measured brain activity during improvisation in classical and jazz pianists, and found higher functional connectivity between prefrontal and premotor areas in the experienced improvisers, which could reflect a more efficient exchange of information within associative networks of importance for generation of music.

The high domain specificity of many of the effects discussed above would make it plausible that causal effects of deliberate practice are involved in the observed expert–novice differences. However, the literature essentially consists of correlational data and observational longitudinal designs, and it is therefore also possible that other causal mechanisms, such as genetic pleiotropy, are involved.

**Expertise and Transfer**

**Near transfer.** Transfer refers to influences of training on nontrained tasks and skills (Adams, 1987; Barnett & Ceci, 2002; R. A. Schmidt & Lee, 1999). One of the most influential theoretical accounts of transfer, the *identical elements theory*, originates with three papers by Thorndike and Woodworth from 1901. These articles discuss results from a series of experiments in which participants were trained on a variety of tasks that included detecting and marking words with particular properties in a text (Thorndike & Woodworth, 1901b), and judgment tasks in which the size of geometrical figures (Thorndike & Woodworth, 1901a; Thorndike & Woodworth, 1901) and weights of objects (Thorndike & Woodworth, 1901a) should be estimated. Participants improved on the trained tasks, but transfer effects to highly similar control tasks (marking words with other properties than those trained; estimating the size of figures with new shapes or sizes, or weights of novel objects) appeared to be remarkably small or negligible. However, the authors found indications of some transfer, in particular when the trained task and the transfer task shared elements (e.g., marking words that contain “s” and “p” at transfer, after training to mark words that contain “s” and “e”). The overall conclusions were that transfer of skills essentially occurs when the trained skill and the transfer task contain identical elements, and—in more general terms—that the mind is “a machine for making particular reactions to particular situations,” which “works in great detail, adapting itself to the special data of which it has had experience” (Thorndike & Woodworth, 1901, pp. 249–250).

Many later studies on both sensorimotor and cognitive skills have confirmed that transfer is limited, and strongly dependent on the amount of shared material between training and transfer task (for reviews, see Adams, 1987; R. A. Schmidt & Lee, 2005; Singley & Anderson, 1989; VanLehn, 1996). For example, Bovair, Kieras, and Polson used a modern operationalization of the identical elements theory—in which procedural knowledge was quantified using a production system model (cf. Singley & Anderson, 1989)—to successfully predict transfer effects as a function of preexisting knowledge in participants learning to handle text-editing software (Bovair, Kieras, & Polson, 1990; Kieras & Bovair, 1986). In line with the identical elements theory, there is also a large body of literature demonstrating part–whole transfer, in which a complex skill is mastered by training its components in isolation (Adams, 1987; R. A. Schmidt & Lee, 2005), and intermanual transfer, in which a sensorimotor skill trained with one hand is rapidly learned with the opposite hand (Hauert, Deiber, & Thut, 2002).

The psychological and neurobiological literature reviewed in the previous sections suggests that the acquisition of expertise is a process of specialization, in which an individual invests time, effort, and neural resources in order to optimize the performance of a limited set of tasks. This implies that there will be few shared elements—to use Thorndike’s terminology—between expert performance and other skills, as the specific adaptations that allow for extremely efficient domain-specific information processing will often be useless in tasks outside the domain of expertise. A reasonable prediction is thus that transfer of expertise to non-domain-relevant tasks should be small.

This prediction is confirmed by empirical data. In fact, even minor changes in the materials or structure of a task can lead to dramatic performance drops, pointing to a specialization of the specialized, that is, differentiation of experts within domains. In Ericsson and colleagues’ case study of memory-span training (Ericsson et al., 1980), the participant SF had a digit memory span of close to 80 digits at the end of training, but virtually no increase in letter memory span. Bilalić, McLeod, and Gobet (2009) studied subgroups of chess experts who had specialized in certain chess openings and compared their performance in memory and problem solving tasks, based on chess positions taken either from the openings they had specialized in or other openings. Large reductions (~1 SD) in performance were found in trials based on the
chess openings that were less familiar to the participants. As already summarized, large performance differences between the perceptual processing of familiar and unfamiliar patterns have been reported for many forms of expertise, including chess, bridge, and music.

Nevertheless, several studies suggest that expertise may show a certain amount of near transfer to similar tasks. Strong chess players have a somewhat better memory for random chess positions than weaker players, although the effect size is much lower than for positions from real games: Gobet and Simon (1996b) found that an increase in chess skill of 400 Elo points yielded an increase in memory capacity of about five pieces for real game positions compared with only about one piece for random positions. Expert athletes may show superior performance in related sports (Rosalie & Müller, 2012, 2014). Such findings are certainly consistent with the possibility of near transfer, which, in part, could be a consequence of shared elements between the studied tasks. However, many studies reporting near transfer rely on cross-sectional data. Thus, it is possible that correlations between expert performance and performance on nontrained tasks also depend on basic abilities that influence performance in both tasks, rather than narrow transfer (i.e., causal effects of training). In line with this possibility, Mosing, Madison, Pedersen, Kuja-Halkola, and Ullén (2014) found that the correlation between musical practice and musical perceptual ability was essentially caused by genetic pleiotropy rather than causal influences of long-term musical training on musical discrimination.

Far transfer. The possibility of far transfer to tasks that are different from the trained task is obviously of larger practical interest, especially in education. If long-term deliberate practice would improve attention and WM, or even general intelligence, one could conceivably also see broad transfer, that is, far transfer to a wide range of cognitive tasks and general scholastic achievement. However, the empirical support for broad transfer of expertise is weak. This could be seen as unsurprising, given that even near-transfer effects are small and that the neural underpinnings of expertise and general intelligence appear to be very different, the latter apparently being more related to global brain properties such as, for example, total gray and white matter volume (Deary, Penke, & Johnson, 2010; McDaniel, 2005; Rushton & Ankeny, 2009), cortical thickness (Karama et al., 2009), and general topological properties of white matter connections (Li et al., 2009). Here, we will briefly summarize findings from four domains that have been extensively studied in this context: music, chess, video games, and special "brain games" designed to improve WM.

The literature on music and cognitive abilities is extensive and has recently been the subject of a comprehensive review by Schellenberg and Weiss (2013). Many studies report positive associations between long-term musical practice and performance on a wide range of cognitive tasks (musical, verbal, mathematical), including tests of general IQ. Correlations between amount of musical training and intelligence are typically small to modest \( r = .15 \) to \( .35 \); Mosing, Madison, Pedersen, & Ullén, 2015; Schellenberg & Weiss, 2013). The large majority of studies use cross-sectional or observational longitudinal designs, and thus are ambiguous with respect to the causal nature of the association. The correlation between musical practice and IQ was recently analyzed in a large Swedish twin cohort (Mosing et al., 2015). When controlling for all genetic and shared environmental influences in a monozygotic intrapair difference design, the association \( r = .15 \) disappeared, and bivariate classical twin modeling indicated that the correlation was largely driven by genetic pleiotropy. These results speak strongly against a causal influence of musical practice on IQ, and highlight that genetic factors are important to consider in any study finding associations between practice and intelligence. Some randomized and quasi-randomized intervention studies have found indications of causal effects of musical training on cognitive abilities (Moreno et al., 2011; Schellenberg, 2004). However, other interventions—including one with no less than 3 years of piano training (Costa-Giomi, 1999)—have found no evidence for far transfer (Bilhartz et al., 1999; Mehr, Schachner, Katz, & Spelke, 2013; Moreno et al., 2009).

Studies of cognitive benefits of chess training have likewise yielded mixed results (Bart, 2014; Gliga & Flesner, 2014; Gobet & Campitelli, 2006). As discussed, superior processing of visual information in chess players is highly limited to ecological stimuli. Ullén, Hambrick, and Mosing (2014) found that the correlation between musical practice and IQ was recently analyzed in a large Swedish twin cohort (Mosing et al., 2015). When controlling for all genetic and shared environmental influences in a monozygotic intrapair difference design, the association \( r = .15 \) disappeared, and bivariate classical twin modeling indicated that the correlation was largely driven by genetic pleiotropy. These results speak strongly against a causal influence of musical practice on IQ, and highlight that genetic factors are important to consider in any study finding associations between practice and intelligence. Some randomized and quasi-randomized intervention studies have found indications of causal effects of musical training on cognitive abilities (Moreno et al., 2011; Schellenberg, 2004). However, other interventions—including one with no less than 3 years of piano training (Costa-Giomi, 1999)—have found no evidence for far transfer (Bilhartz et al., 1999; Mehr, Schachner, Katz, & Spelke, 2013; Moreno et al., 2009).

The literature on video game expertise and cognitive abilities is relatively large (for reviews, see Boot, Blakely, & Simons, 2011; Ferguson, 2007; Latham, Patston, & Tippett, 2013). Many of these studies have focused on possible transfer effects on visual tasks, such as reaction time (RT), visuospatial rotation, visual search, and visual attention tasks. Cross-sectional studies comparing expert video gamers with novices have commonly found superior performance of gamers on such tasks (see, e.g., Bialystok, 2006; Castel, Pratt, & Drummond, 2005; Greenfield, DeWinstanley, Kilpatrick, & Kaye, 1994; West, Stevens, Pun, & Pratt, 2008; Yuji, 1996), and Ferguson (2007) found a meta-analytic pooled correlation of \( r = .36 \) between violent video game playing and measures of visuospatial cognition. As for chess and music, results of longitudinal studies are inconsistent, with both positive findings and null effects of interventions (Boot et al., 2011; Latham et al., 2013), and the meta-analysis by Ferguson indicates the presence of a publication bias. One can note that Sims and Mayer (2002), in a carefully designed study, found—entirely in line with predictions from expertise research—that skilled Tetris players outperform non-players on visuospatial rotation tasks that involve Tetris pieces, or highly similar visual shapes, but not on other spatial tests (Experiment 1); furthermore, no transfer effects of Tetris training on spatial ability were found in a longitudinal follow-up study (Experiment 2). In a recent study Gobet et al. (2014) found no effects of video game expertise on an attention task (the Eriksen Flanker Compatibility Task), nor on a visual change detection task.

"Brain games" designed to increase WM capacity typically include extensive practice of various WM tasks in which the difficulty level changes adaptively during training. Many large-scale studies demonstrate that such training improves performance on the trained tasks (Klingberg, 2010; Shinaver, Entwistle, & Söderqvist, 2014), but the extent to which the training also gives far transfer to other contexts is highly controversial. A detailed discussion of these disagreements is beyond the scope of the present review, but we note that several recent reviews and meta-analyses have found little evidence for broad transfer of WM training (Melby-Lervåg & Hulme, 2013; Rapport, Orban, Kofler, & Friedman, 2013; Shipstead, Hicks, & Engle, 2012). However, for a different interpretation of the literature, see, for example,
Shinaver et al. (2014). In any case, it appears evident that the acquisition of relatively narrow task-specific expertise is one important mechanism underlying performance improvement in WM training tasks, as in other forms of skill learning (Ericsson et al., 1980; Shipstead, Redick, & Engle, 2012). In line with this, performance in visuospatial WM tasks depends not just on WM capacity as typically operationalized (i.e., the number of stimuli to be remembered) but also on the spatial pattern of the stimuli (Nutley, Söderqvist, Bryde, Humphreys, & Klingberg, 2010) and their familiarity, i.e., previous exposure to the same stimuli (Moore, Cohen, & Ranganath, 2006; Zimmer, Popp, Reith, & Krick, 2012).

Challenges to Deliberate Practice Theory

The literature summarized in the previous section illustrates that research on expert performance and its relation to deliberate practice has advanced scientific understanding of expertise and skill learning in important ways. Nevertheless, recent research has seriously challenged the central tenets of the expert performance framework. In short, deliberate practice theory appears incompatible with the findings of a growing number of studies, which show that variables other than deliberate practice are highly relevant to expert performance. We will now turn to these data.

Practice, Traits, and Expertise: Phenotypic Studies

How much variance in expert performance is actually explained by deliberate practice? An obvious prediction of deliberate practice theory is that individual differences in deliberate practice should be highly correlated with measures of expert performance. Indeed, Ericsson et al. (1993) claimed that “individual differences in ultimate performance can largely be accounted for by differential amounts of past and current levels of practice” (p. 392, italics added). Because the theory actually implies that, with the exception of certain sports in which height and body size play an obvious role, deliberate practice is the only variable of importance for expert performance (Ericsson et al., 1993), it would be tempting to conclude that we should predict not just a high, but a more or less perfect, correlation. However, in reality, neither deliberate practice nor expert performance can be measured with perfect reliability. Nevertheless, even a liberal reading of the deliberate practice theory leads to the minimal prediction that most of the variance in expert performance should be explained by deliberate practice, taking measurement error into account.

The available evidence does not support this prediction. A number of recent reviews and meta-analyses have provided estimates of the proportion of variance in expert performance explained by deliberate practice in different domains (Hambrick, Oswald, et al., 2014; Macnamara et al., 2014; Platz et al., 2014). The estimates range from less than 1% for professional expertise in Macnamara et al. (2014), to 36% for music in Platz et al. (2014). Another important finding was that correlations between deliberate practice and expert performance varied systematically with certain moderator variables as well as between domains (Macnamara et al., 2014). Expert performance in games and music showed the strongest relation to practice. Predictable tasks showed a relatively strong relation to deliberate practice, whereas more unpredictable tasks showed a weaker relationship. Furthermore, in striking contrast to predictions from the expert performance framework (Ericsson et al., 1993; Ericsson & Lehmann, 1996), relations were stronger for retrospective estimates of deliberate practice than for deliberate practice logs, which presumably yield more valid estimates of deliberate practice, and when professional status or ranking was used as an index of expertise than when experimental measures of expert performance were used (Macnamara et al., 2014).

This evidence converges on two conclusions. First, relations between retrospective practice estimates and expertise may be inflated, for example, because participants’ perceptions of their current ability and achievements influence their estimates of earlier practice. Second, experimental measures of—inevitably limited—aspects of expert performance may not be more valid measures than ratings or group membership, which could be based on a more global and balanced assessment of the totality of skills needed in a domain. Taken together, these observations strongly suggest that expert performance depends also on other factors than deliberate practice. In the next section, we review the literature on what some of these factors may be.

Expert performance and physical traits. Even the staunchest proponents of deliberate practice have consistently allowed for one exception to the deliberate practice theory, namely, that in certain sports, the physical traits of body size and height may influence expert performance over and above deliberate practice (Ericsson, 2007, 2014c; Ericsson et al., 1993). The importance of height for sports like high jumping and basketball is indeed self-evident. In reality, however, there are many anthropometric variables that influence biomechanical body properties of relevance for physical expert performance (Epstein, 2013). Energy economy during running is related to body height, as well as ponderal index, physique, and morphology of the feet, the legs, and the pelvis (Anderson, 1996). Runners tend to have longer legs than swimmers, and the size of the upper extremities matters for sports like handball. For example, Debanne and Laffaye (2011) found that ball throwing velocity is substantially correlated with both general anthropometric properties such as height and body mass ($r = .55$ to $.70$), and handball-specific anthropometric properties such as arm span and hand size ($r = .35$ to $.51$).

Anthropometric properties are also of importance for other expert performance that involves physical activity, such as music performance and dancing. Dancers specializing in dances with different functional requirements show characteristic group differences in height and somatotype (Livi et al., 2013). Hand morphology influences dexterity in pianists, and a small hand makes it difficult to play wide chords and increases the risk for overuse disorders (Sakai & Shimawaki, 2010; Wagner, 1988).

Critically from the standpoint of evaluating deliberate practice theory, anthropometric variables are strongly influenced by genetic factors, with heritabilities typically in the range 70% to 90% for measures like body height (Silventoinen, Magnusson, Tylenius, Kaprio, & Rasmussen, 2008; Silventoinen et al., 2003), arm and finger lengths (Poveda, Jelenkovic, Susanne, & Rebato, 2010), length of the eye (Kim et al., 2013; Lyhne, Sjølie, Kyvik, & Green, 2001), and head size (Smit et al., 2010). In principle, a high heritability does not imply that a trait cannot be modified by environmental interventions, but we are aware of no empirical demonstrations that deliberate practice has significant effects on
physical traits that reflect the dimensions of hard tissue, such as height or the length and proportions of the extremities.

**Expert performance and cognitive ability.** The claim that cognitive abilities do not limit expert performance is a particularly controversial claim of deliberate practice theory, given the strong theoretical reasons to expect the opposite. First, one of the most well-established findings in psychology is that the general factor of intelligence—psychometric $g$—correlates with learning, especially in cognitively demanding tasks (Ackerman, 1996; Ackerman & Ciancio, 2000; Jensen, 1998). Deliberate practice is, by definition, explicit and characterized by a conscious search for strategies to improve performance (Ericsson, 2007). In other words, deliberate practice requires sustained attention and metacognition, both of which are correlated with intelligence (Schweizer & Moosbrugger, 2004; Schweizer, Moosbrugger, & Goldhammer, 2005; Stankov, 2000). Regardless of the mechanisms underlying this correlation, one might thus expect that engagement in deliberate practice is more correlated with intelligence than less deliberate forms of practice.

Second, many forms of expert performance are intrinsically unpredictable and involve online processing of novel information, even at high levels of expertise. Such tasks are presumably difficult to optimize through the processes of automation, chunking, and efficient WM-LTM interactions, discussed previously, and could, for that reason, be highly dependent on intelligence and executive control. Finally, intelligence shows weak to moderate correlations with RT ($r = -.3$ to $-.5$; Deary, Der, & Ford, 2001; Jensen, 2006), processing speed ($r = .3$ to $.5$; Neubauer & Knorr, 1998; Rindermann & Neubauer, 2004), and temporal and nontemporal sensory discrimination ($r = .2$ to $.5$; Spearman, C. (1904); Rammsayer & Brandler, 2007; Troche & Rammsayer, 2009)—basic capacities that appear likely to be important for expert performance in many domains.

Consistent with these observations, growing evidence indicates that expert performance is indeed related to cognitive ability. A large literature shows that intelligence predicts educational achievement, an outcome that is obviously related to the acquisition of expertise (e.g., in reading and arithmetic). Correlations between educational achievement and IQ typically fall in the range $r = .4$ to $.7$ (Deary, Strand, Smith, & Fernandes, 2007; Krapohl et al., 2014). Intelligence is also correlated with occupation and performance in the workplace. The mean IQ of employees is higher in occupations with higher demands on complex information processing and cognitive skills (Gottfredson, 2003; Hunt, 2011), and a meta-analysis by F. L. Schmidt and Hunter (1998) found general mental ability to predict individual differences in job performance, with a mean validity of .63. Furthermore, as illustrated by findings from the Study of Mathematically Precocious Youth project (Lubinski & Benbow, 2006), expert performance and achievement are predicted by differences in cognitive ability, even within the highest percentile. Recent findings from this research show that, for example, participants who scored at or above the 99.99th percentile on the verbal or mathematical SAT before Age 13 display outstanding achievements as adults within a broad range of expertise domains that include science, technology, engineering and mathematics (STEM) fields, as well as the arts, business, and law (Kell, Lubinski, & Benbow, 2013). The domains of expertise vary with ability profile, so that high scorers on the verbal SAT tend to excel within arts, humanities, and writing, whereas high scorers on the mathematical SAT display top achievements within STEM fields (Kell et al., 2013).

In contrast, studies on the relation between study time and grade point average have found surprisingly weak effects (Plant, Ericsson, Hill, & Asberg, 2005), and the meta-analysis of Macnamara et al. (2014) found that deliberate practice, on average, predicted only 4% of the variance in expert performance in educational contexts.

The contributions of intelligence and deliberate practice to expert performance have also been assessed in studies that include more direct measures of expert performance or skill level. Null effects of cognitive abilities on expert performance in experts have been reported for predictable tasks such as typing (Keith & Ericsson, 2007). In contrast, IQ predicts performance in chess experts: Grabner, Stern, and Neubauer (2007) found a correlation of $r = .35$ between general intelligence and Elo rating in Austrian tournament chess players. Meinz and Hambrick (2010) found a partial correlation of $r = .37$ between WM capacity and sight-reading ability in musicians, after controlling for deliberate practice. In line with these findings, Ackerman (1987) has argued that the associations between cognitive abilities and performance, at different stages of training, should depend strongly on how demanding the task is in terms of attentional resources, and whether it is consistent or predictable enough to be efficiently automated during training. For attention-demanding tasks that are difficult to automate, this framework predicts that significant correlations between cognitive ability and performance should remain after practice; support for this was indeed found in reanalyses of several classical datasets from Thordike, Woodrow, Fleishman, and others (Ackerman, 1987). Later studies have corroborated that correlations between performance and intelligence remain and may even increase during several hours of training of complex, low-predictability tasks (Ackerman, 2014; Ackerman & Ciancio, 2000).

**Expert performance and personality.** There are several reasons to expect relations between expert performance and personality. First, personality—which is moderately related to vocational interests (Ackerman & Heggestad, 1997; Harris, Vernon, Johnson, & Jang, 2006; Holland, 1997; McKay & Tokar, 2012)—is likely to influence which domain an individual chooses to become an expert in. Second, personality could impact both amount and quality of practice. Third, personality may even have a direct influence on expert performance, independent of deliberate practice, in certain domains. For example, extraversion could influence expert performance in stage arts, and personality traits related to emotional competence could play a role for expert performance in domains that involve a high degree of interpersonal interaction (e.g., psychotherapy).

The few relevant studies in the literature suggest that personality is indeed related to deliberate practice. Grit, a personality variable reflecting a persistence in pursuing long-term goals that is strongly ($r = .77$) correlated with conscientiousness (Duckworth, Peterson, Matthews, & Kelly, 2007), predicted deliberate practice ($r = .30$) and performance ($r = .17$) in a spelling competition for children, whereas openness did not (Duckworth, Kirby, Tsukayama, Bernstein, & Ericsson, 2011). In contrast, Butkovic, Ullén, and Mosing (2015) found voluntary music practice to be predicted by openness ($r = .31$) and a proneness to experience psychological flow during musical activities ($r = .46$; see also McPherson & McCormick, 1999; Smith, 2005). In a study of chess players, Grabner et al.
(2007) found associations between playing strength and personality measures reflecting domain-specific performance motivation and emotion expression control ($r = .27$ to $.39$). As another example, Miksza (2011) found impulsivity in musicians to be related to less-well-organized practice and poorer achievement with correlations in the range $r = - .2$ to $- .4$ between impulsivity and practicing behaviors involving repetition.

There is also evidence that personality influences expert performance, above and beyond deliberate practice. In particular, large-scale studies by Ackerman and colleagues have demonstrated how correlations between vocational interests, personality, and abilities form different “trait complexes,” with different personality profiles of individuals active in different domains (social, scientific, clerical, cultural; Ackerman, 1996; Ackerman, Chamorro-Premuzic, & Furnham, 2011; Ackerman & Heggestad, 1997).

Genetic influences on deliberate practice, expert performance, and associated traits. Given that most complex traits, including those associated with expert performance, such as personality, interests, and cognitive abilities, have been shown to be partly heritable, it is reasonable to also expect some genetic influences on variation in expert performance (Bouchard & McGue, 2003; Harris et al., 2006; Poldermann et al., 2015; Plomin & Spinath, 2004). Moreover, as summarized earlier, expertise critically depends on neural plasticity. Experts and nonexperts show extensive differences in regional brain anatomy in both the gray and the white matter. Animal data supports that macroanatomical effects of long-term training in part reflect formation of new neuronal processes and synapses (see, e.g., Kleim et al., 2004). Functional reorganizations in experts include both altered functional properties of specific brain regions and changes in connectivity between regions—adaptations that presumably serve to optimize the processing of domain-specific information.

The animal literature is replete with evidence for the importance of genetic factors for neural plasticity, many of which appear to be highly conserved through the animal kingdom (Frank & Fossella, 2011). In light of this, it would be most surprising to find that genetic factors played no role for the acquisition of expert performance. Indeed, recent twin studies have confirmed substantial genetic effects on music practice (Hambrick & Tucker-Drob, 2015; Mosing, Madison, et al., 2014) as well as self-rated expertise (Vinkhuizen, van der Sluis, Posthuma, & Boomsma, 2009). Furthermore, genetic pleiotropy is a major mechanism behind associations between openness, psychological flow experiences during musical activities, and music practice (Butkovic et al., 2015), as well as between intelligence and music practice (Mosing et al., 2015).

Behavior genetic studies have also shed light on the nature of phenotypic associations between practice and performance. Musical aptitude is influenced by genetic factors (Drayna, Manichaikul, de Lange, Snieder, & Spector, 2001; Oikkonen et al., 2015; Ukkola-Vuoti et al., 2013), and Mosing, Madison, et al. (2014) found genetic pleiotropy to be the major factor behind associations between music practice and musical aptitude. This was supported by two analyses. First, the results of a monozygotic intrapair difference (cotwin control) design showed that the phenotypic relation between practice and ability ($r = .15$) disappeared entirely when controlling for both genetic factors and shared environment. Furthermore, bivariate classical twin modeling suggested that the association was essentially driven by common genes, that is, pleiotropy. These results thus speak strongly against a causal effect of practice on musical perceptual ability. Similarly, Ullén, Mosing, and Madison (2015) recently reported that the association between music practice and accuracy of motor timing disappears when controlling for genetics and shared environment. Furthermore, Hambrick and Tucker-Drob (2015) reported common genetic effects on musical practice and achievement, and found indications that the importance of genetic factors for musical achievement—contrary to predictions from the deliberate practice theory—increased with practice. This is line with a twin study on training of the rotary pursuit task, which found genetic influences on performance as well as rate of learning, with an increased heritability of performance after three days of training (Fox, Hershberger, & Bouchard, 1996).

Rethinking Expertise

Deliberate Practice Theory and Recent Empirical Challenges

Criticisms of the extreme claims of deliberate practice theory are not new and have, on several occasions, been highlighted in published exchanges focusing on various aspects of the expert performance framework, including the question of innate talent (see Howe et al., 1998, with commentaries), the relative importance of deliberate practice and other variables for expert performance (Dettman, 2014; Hambrick, Altmann, Oswald, Meinz, & Gobet, 2014; Hambrick, Oswald, et al., 2014; Platz et al., 2014), and giftedness (see Ericsson et al., 2007a, with commentaries). We would like to emphasize that—as pointed out by philosopher of science Imre Lakatos (1978)—it is essential to evaluate a model in terms of how well it can explain major findings in a field, taken together, and how fruitful it is in terms of generating new hypotheses. In any complex field, unpredicted results of individual experiments can usually easily be made compatible with a theory, using one ad hoc explanation or another (Lakatos, 1978). We believe, therefore, that it is instructive and revealing to briefly examine a few of the more common themes in recent debates on deliberate practice theory, keeping these observations in mind.

First, as detailed in the previous section, meta-analyses demonstrate that deliberate practice fails to account for all, nearly all, or even most of the variance in expert performance, and often even explains only a surprisingly small proportion of the total variance. A common objection to such findings is to question the study’s validity, and more specifically, to suggest that the participants did not engage in sufficiently deliberate practice that and/or the employed practice measure did not tap deliberate practice unambiguously enough (Ericsson, 2014b, 2014c; Ericsson, Roring, & Nandgopal, 2007b). This argument may appear circular—deliberate practice explains expert performance, and if the data show something else, the practice was not real deliberate practice—but of course the point would be completely valid if there were a general agreement on exactly which activities qualify as deliberate practice in various domains, as well as solid data showing that deliberate practice, with this definition, predicts expert performance, whereas other forms of practice do not. However, the distinction between deliberate practice and other forms of practice is in reality often very unclear (Howard, 2009). Furthermore, there is abundant ev-
idence showing that forms of domain-relevant experience other than deliberate practice lead to improvements of performance, and that the optimal practice style can vary depending on personality (Howard, 2009, 2012; Kluge, Ritzmann, Burkolter, & Sauer, 2011). In fact, entirely implicit forms of learning from experience and exposure are important, for example, for the acquisition of linguistic expertise (Forkstam & Petersson, 2005; Perruchet & Pacton, 2006). Finally, there is evidence from chess that the number of games played—a domain-related activity that cannot be regarded as solitary deliberate practice in any sense of the term—is a strong predictor of skill level (Howard, 2012).

A second common objection is that studies showing modest effects of deliberate practice have ignored effects of injury, disease, and related variables (Ericsson, 2014c). These variables should obviously be considered among the many variables, apart from deliberate practice, that potentially influence expert performance. However, it is important to note that the proportion of variance in expert performance that should be explained by deliberate practice, taking the mentioned variables into account, is not specified by deliberate practice theory. This makes the proposal impossible to evaluate empirically (i.e., to falsify). More importantly, deliberate practice theory provides no plausible explanation for recent empirical findings on this issue, including the systematic differences in deliberate-practice/expert-performance relations between tasks and domains, and their dependence on moderator variables. Deliberate practice theory does not explain why, for example, deliberate practice explains a relatively large proportion of expert performance variance in musicians and athletes—groups highly susceptible to injury and overuse—and a much smaller amount of the variance in domains in which injury and overuse are uncommon, such as professions with minimal physical demands (e.g., sales; see Sonnentag & Kleine, 2000).

A third recurring controversy is whether tasks that show less impressive correlations with deliberate practice and/or substantial relations to variables other than deliberate practice qualify as true forms of expert performance. One example is musical aptitude, operationalized as the ability to discriminate musical melodies, rhythms, and pitches. This ability is related to intelligence (Lynn & Gault, 1986; Lynn, Wilson, & Gault, 1989) and music practice (Ullén, Mosing, Holm, Eriksson, & Madison, 2014). As already discussed, it has recently been found that these associations are essentially driven by common genetic influences (Mosing, Madsen, et al., 2014; Mosing, Pedersen, Madison, & Ullén, 2014). In a recent commentary, Ericsson (2014c) dismisses studies of music discrimination because they “test memory and perception of music rather than music performance” (p. 85). However, the capability to discriminate musical sounds is essential for music performance. Furthermore, musical discrimination is commonly measured in entrance exams to music colleges, and correlates nontrivially with music practice, musician status, and musical achievement (Seashore, 1938; Ullén et al., 2014; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). One can note that in other contexts, when expert performance does correlate with deliberate practice but not with professional status, this has been taken not as a challenge to the expert performance paradigm or as evidence that the expert performance task is unrepresentative, but rather as support for that professional status is an unreliable and potentially irrelevant outcome in expertise research (Ericsson, 2014c; Ericsson & Lehmann, 1996).

A closely related issue is the question of who does and does not qualify as an expert in a field. Early studies suggested a fairly liberal criterion for expertise, namely, that anyone performing at two standard deviations or more above the mean of the general population on a representative task is an expert (Ericsson & Charness, 1994). This definition regards expertise as one end of a continuous distribution and would—reasonably—imply that many individuals should be considered as experts in various fields. However, as discussed, an important additional characteristic of expertise is that it requires relatively extended periods of training (months or years). An often quoted rule of thumb is that 10 years of practice might be required for “mastery” of a domain (Ericsson & Charness, 1994; Simon & Chase, 1973). Recently, Ericsson (2014c) has gone even further, suggesting that true expertise may only be displayed by “less than a handful of individuals” (p. 100). Unfortunately, these unclear and arbitrary distinctions between more mundane forms of expertise, mastery, and eminence have been utilized to circumvent challenges from studies demonstrating effects of nondeliberate practice variables on expert performance. For instance, Ericsson considers a large-scale study on the top 5% of 8th grade readers—which demonstrates genetic influences on reading expertise (Plomin, Shakshaft, McMillan, & Trzaskowski, 2014)—as irrelevant, as the studied behavior would not qualify as expert performance. Ericsson even suggests that twin studies are in principle uninformative for expertise research, as too few twins display true expertise to perform behavioral genetic analyses. However, clearly, many twins display expertise by the two-standard-deviation criterion.

Fourth, there is the striking observation of secular increases in top levels of expert performance in certain domains (e.g., sports and chess) during the last decades (Lehmann, 1998). This has sometimes been brought forward as evidence that genetic factors are unimportant for expert performance, because it would appear unlikely that significant changes in relevant gene frequencies could have taken place during such a short time period (Ericsson et al., 2007a). However, in general, secular changes in mean level of a trait do not imply that the population variance at a given point in time is nongenetic in nature; well-known examples of highly heritable traits that have shown dramatic increases in Western countries during the last century include body height (Silventoinen et al., 2003) and IQ (Plomin & Spinath, 2004).

Finally, we would like to consider one of the most common, but also most scientifically problematic, arguments in defense of deliberate practice theory. This is the argument that for deliberate practice theory to be falsified it would have to be demonstrated that variables other than deliberate practice—including traits like intelligence, or genetic factors—limit the ultimate level of expert performance that an individual hypothetically could achieve (Ericsson et al., 2007b). The problem with this proposal is that this could hardly be accomplished, even in principle, because it would require the investigation of an almost infinite number of counterfactual scenarios. Science can investigate existing phenomena in the world as we find it, but it can rarely prove that something never could take place (see, e.g., Plomin et al., 2014). If this suggestion is taken seriously, it would be doubtful whether deliberate practice theory is a falsifiable theory that can be tested empirically. In contrast, a reasonable interpretation of deliberate practice theory does give rise to the testable prediction that most of the variance in expert performance is explained by deliberate practice or, stated
differently, that the influence of nondeliberate practice variables on the probability of acquiring expertise is small or insignificant (see the previous section).

A Multifactorial Gene–Environment Interaction Framework for Expertise Studies

Deliberate practice theory posits that only one variable, deliberate practice, is of key importance for expert performance in all domains of expertise. However, the unavoidable conclusion appears to be that this model is too simple: As discussed in detail in the previous sections, deliberate practice theory fails to account for several key findings relating to expertise and expert performance. Accordingly, we believe it is essential for progress in the field that researchers turn to a more progressive, multifactorial framework for expertise studies that allows for a systematic investigation of all factors that influence expert performance and that can give rise to novel, testable hypotheses. An important part of the motivation for a paradigm shift is methodological. If a researcher works on the assumption that deliberate practice is all that matters for expert performance, he or she is likely to choose experimental designs that allow for a characterization of practice effects, but fail to consider the importance of other variables, such as cognitive abilities. In fact, a striking feature of the field as a whole is that relatively few studies have studied expertise in genetically informative samples or investigated the relative importance of deliberate practice and other phenotypic variables for expertise in a systematic manner. The relatively few recent studies that have addressed such issues have indeed provided important evidence for the involvement of variables other than deliberate practice in expertise (see the previous section), and it appears likely that continued research along these lines will be very fruitful for our understanding of expert performance. In this final section, we will summarize the general features of a possible alternative framework for expertise studies, which we dub the *multifactorial gene–environment interaction model* (the MGIM). Key features of this framework are schematically illustrated in Figure 1. Variables given in italics under each broader category are specific examples of factors that have been shown empirically to be involved in expertise (see earlier sections).

Some general similarities between the MGIM and earlier theories of intellectual development should be mentioned. The potential roles of intelligence and interactions between intelligence and practice for the development of expert performance in the MGIM could be seen as generally similar to Cattell’s (1987) model of the development of specialized knowledge in his *investment theory*. An even closer relationship can be seen with Ackerman’s (1996)
PPIK theory, in which adult intellectual development is modeled as a function of intelligence as process, personality, interests, and intelligence as knowledge. The MGIM focuses on the development of expert performance, rather than general intellectual development, but is similarly multifactorial and emphasizes that a wide range of both psychological and physical individual difference variables are likely to be important for expert skills, in a domain-specific way (Figure 1, upper part). A second important feature of the MGIM is that the development of expertise is regarded from a gene–environment interaction perspective, that is, genetic factors, nongenetic factors, and their interactions are considered as important influences on expert performance, deliberate practice, expertise-relevant traits, and their covariation (Figure 1, lower part).

The role of deliberate practice in the MGIM and deliberate practice theory. It may be useful and prevent misunderstandings to begin with the points at which the MGIM and deliberate practice theory agree. One major point of agreement is that deliberate practice is important for expertise. There is ample empirical evidence that deliberate practice is an important predictor of expert performance, and in many domains, it may well turn out to be the most important single predictor, even when other variables are taken into account (Macnamara et al., 2014). The models also agree that one important reason for observed deliberate-practice/expert-performance correlations is that deliberate practice has causal influences on expert performance. Practice may not make perfect, but it certainly helps. As detailed earlier, prolonged deliberate practice over months and years may lead to extensive anatomical and functional reorganizations of brain circuits that in turn underpin the often astonishing levels of domain-specific performance we see in experts. As discussed earlier, these reorganizations involve brain circuits that enable rapid and efficient interactions between LTM and WM (LTM–WM), as well as sensorimotor skills and optimized processing of domain-specific sensory information; deliberate practice can also modify some physical body properties, such as muscle strength, that are of importance for certain forms of expertise (see Figure 1). Deliberate practice theory and the MGIM agree in that expert performance is the consequence of specialization rather than improvements of general capacity; transfer effects of deliberate practice to nontrained tasks are often negligible.

This stated, the two models also differ in their view of deliberate practice in several important ways. The MGIM emphasizes that deliberate practice cannot be seen as an entirely environmental phenomenon: Engagement in deliberate practice is not a way to “circumvent” genetic limitations. Rather, behavior genetic studies strongly suggest that the quantity, and presumably also quality, of deliberate practice is itself subject to substantial genetic influences. Although deliberate practice certainly has causal influences on expert performance, recent data on musical expert performance show that reverse causality and genetic pleiotropy should be taken as serious empirical possibilities: For both musical discrimination ability (Mosing, Madison, et al., 2014) and accuracy of motor timing (Ullén et al., 2015), it has recently been demonstrated that associations with musical practice are driven essentially by common genetic factors rather than causal effects of practice.

Finally, an important feature of the MGIM is that deliberate practice can be influenced in various ways by variables (abilities, personality, interests, motivation) that potentially impact which domain an individual elects to invest time in, as well as the intensity and quality of the practice itself. Figure 1 provides specific empirically supported examples of such influences (see earlier sections): Both grit and openness have been shown to be associated with the amount of deliberate practice in different domains (Butkovic et al., 2015; Duckworth et al., 2011). For the domain of music, impulsivity has been shown to influence practice quality (Miksza, 2006, 2011). For a general discussion of relations between interests, personality, abilities, and domain of expertise, see, for example, Ackerman (1996), Ackerman et al. (2011), and Kell et al. (2013).

The MGIM at the phenotypic level. An essential difference between deliberate practice theory and the MGIM is that latter assumes, as a central tenet of the model, that expert performance can be influenced directly by a number of other variables than practice. As discussed in the previous section, it appears important to consider a wide range of individual difference variables for expertise, above and beyond effects of practice. These could potentially include different modalities of psychological individual differences—abilities, personality, interests, social attitudes, motivational variables—as well as physical traits (Figure 1, upper part). It appears inevitable that the importance of individual variables will show considerable variation across domains. As summarized earlier, there is clear evidence that, even at high levels of expertise, intelligence and related constructs such as WM are of importance for domains that require the processing of complex and novel information. This includes musical expertise (Meinz & Hambrick, 2010) and chess (Grabner, 2014), and is further corroborated by the general observation that individual differences in spatial and verbal cognitive abilities predict domain-specific performance and achievement, even at the highest levels of intelligence (Kell et al., 2013). Physical traits will matter primarily for forms of expert performance that involve sensorimotor acts: Height and body size are, for example, important for many sports (Ericsson, 2007; Ericsson et al., 1993), whereas biomechanical properties of the hand are important for instrumental musicians. An important general feature of the MGIM framework is thus that the role of variables other than deliberate practice for expertise may vary between domains.

A second general point is that relations between expert performance and predictor variables are likely to be complex and involve multiplicative as well as additive effects. To give one example, it appears plausible that deliberate practice could interact with intelligence and personality. As discussed previously, deliberate practice would be expected to load strongly on cognitive mechanisms such as attention and metacognition. Not least for cognitively demanding domains, it thus appears possible that the effect of each hour of deliberate practice varies as a function of intelligence—in other words, that there is an interaction between deliberate practice and intelligence. Furthermore, studies on musicians have found relations between personality (impulsivity) and practicing habits (Miksza, 2006, 2011); this suggests the possibility of interactions between deliberate practice and personality.

Such nonlinear relationships are also of interest in relation to the exceptional achievements of prodigies and eminent creators, compared with individuals with comparable levels of deliberate practice. These provide yet another anomaly for deliberate practice theory, whereas it would seem that they can be readily explained within the MGIM if performance is assumed to depend on both
additive effects and multiplicative interactions between deliberate practice and other variables (Lykken, 2005; Mosing & Ullén, in press; Simonton, 1984).

The MGIM: Genetic perspectives. A major reason for the shortcomings of deliberate practice theory is its extreme environmentalist perspective. In contrast, the MGIM assumes—on the basis of both available data and theoretical considerations—that genetic factors are likely to be important for expertise in a number of ways (see Figure 1, lower part). First, one should consider practice-independent genetic effects on traits that are relevant for expert performance. One obvious variable of interest in this context is intelligence, which is highly heritable (Jensen, 1998; Plomin & Spinath, 2004), and likely to be important for many forms of expert performance, even at high levels of deliberate practice (see previous sections). Many genetic effects on anthropometric properties are also likely to fall in this category. Training-independent genetic effects should also be considered in relation to more narrow abilities of relevance for expert performance. One already-mentioned example is musical discrimination, which shows a genetic overlap with intelligence as well as more specific genetic influences (Mosing, Pedersen, et al., 2014).

Other genetic factors will affect expert performance through their effects on deliberate practice and its covariates. This suggests that the genetic architecture of expert performance may be different at different levels of deliberate practice. Indeed, preliminary findings suggest that the heritability of expert performance may actually increase rather than decrease with deliberate practice, contrary to predictions of deliberate practice theory (Hambrick & Tucker-Drob, 2015). More generally, this also highlights the potential importance of more complex interplay between genes and environment—that is, gene–environment interactions (Kennedy et al., 2011; Schellenberg, 2015) and various forms of gene–environment covariation (Scarr, 1996), for expertise.

Conclusion

The last decade has witnessed an impressive development in expertise research through the combination of experimental psychology with more novel methods, in particular, human neuroimaging techniques. An important reason for this is that cognitive modeling of expert performance now can be integrated with neurobiological studies of its underlying mechanisms. We believe there is every reason to expect continued development of the field, if expertise research is performed in a multivariate framework that also includes methods from differential psychology and behavior genetics. This will presumably lead to more complex, but also more interesting and fruitful, models of expertise than currently exist, in which the development of domain-specific competence is seen to depend on genes, environment, practice, and traits in complex interactions.

References


the National Academy of Sciences of the United States of America, 107, 11163–11170. http://dx.doi.org/10.1073/pnas.1005062107

Mehr, S. A., Schachner, A., Katz, R. C., & Spelke, E. S. (2013). Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment. *PLoS ONE, 8*(12), e82007. http://dx.doi.org/10.1371/journal.pone.0082007


Thordike, R. L., & Woodworth, R. S. (1901b). The influence of improvement in one mental function upon the efficiency of other functions. III. Functions involving attention, observation and discrimination. Psychological Review, 8, 553–564. http://dx.doi.org/10.1037/h0071363


Received January 26, 2015
Revision received August 17, 2015
Accepted August 31, 2015