How Important Is Deliberate Practice? Reply to Ericsson (2016)

Brooke N. Macnamara¹, David Z. Hambrick², and David Moreau³
¹Case Western Reserve University, ²Michigan State University, and ³University of Auckland

Abstract

Ericsson and colleagues’ view that individual differences in expertise can largely be accounted for by accumulated deliberate practice is not supported by the available empirical evidence. Extending earlier work (Macnamara, Hambrick, & Oswald, 2014), we found that deliberate practice accounted for a sizeable amount of variance in sports performance (18%), but it left a much larger amount unexplained. Ericsson’s (2016, this issue) evaluation of our research is undercut by contradictions, omissions, and errors. We agree with Ericsson that future longitudinal research will deepen understanding of expertise, but our goal was to evaluate the importance of deliberate practice based on existing evidence.

We thank Anders Ericsson for his commentary on Macnamara, Moreau, and Hambrick (2016). In this meta-analysis, we found that deliberate practice, though important, left a large amount of variance in sports performance unexplained. Thus, we concluded that deliberate practice is not as important as Ericsson and colleagues have argued. In his commentary, Ericsson (2016, this issue) rejects our conclusion, claiming that our methods were flawed. We credit Ericsson for his influential work; however, as before (see Hambrick et al., 2014), his evaluation of our research is undermined by contradictions and inconsistencies in his arguments.

Most notably, Ericsson (2016) claims we deviated from his definition of deliberate practice as training designed by a teacher. Ericsson (1998) stated, “Ericsson, Krampe, and Tesch-Römer (1993) proposed the term deliberate practice to refer to those training activities that were designed solely for the purpose of improving individuals’ performance by a teacher or the performers themselves” (p. 84, emphasis added; see also Keith & Ericsson, 2007). Moreover, in several of his own studies, we could find no record that participants were instructed to restrict deliberate practice estimates to teacher-designed activities (e.g., Duffy, Baluch, & Ericsson, 2004; Ericsson et al., 1993). Accordingly, we allowed that deliberate practice could be designed by a teacher/coach or the performers themselves.

As another example, Ericsson (2016) criticizes us for including studies that used categorical expertise measures (e.g., selection to national team) rather than objective measures, as well as studies that measured group activities. Yet, he has repeatedly used some of these very studies to argue for the importance of deliberate practice (e.g., Helsen, Starkes, & Hodges, 1998; Hodges & Starkes, 1996; see Ericsson, 1998, 2006; Ericsson, Nandagopal, & Roring, 2005; Keith & Ericsson, 2007). Furthermore, compared with the overall average variance explained by deliberate practice (18%; Fig. 1a), results were similar for studies that used objective performance measures (20%; Fig. 1b) and studies that measured individual (non-group) practice (22%; Fig. 1c).

Ericsson (2016) further argues that our approaches would converge more if we analyzed “yearly estimates of hours” (p. 353) of practice. Yet, his major conclusions about the importance of deliberate practice are based on accumulated estimates (e.g., Duffy et al., 2004; Ericsson et al., 1993; Hutchinson, Sachs-Ericsson, & Ericsson, 2013). This focus dictated ours.

Finally, defining an expert musician as someone “who had competed in national competitions with good outcomes” (p. 4), Ericsson (2016) argues that results of Hambrick and Tucker-Drob’s (2014) and Mosing, Madison, Pedersen, Kuja-Halkola, and Ullén’s (2014)
behavioral-genetic studies “cannot be generalized to expert musicians” (p. 352) because the samples included amateurs. Yet, he refers to the skilled bowlers in Harris’ (2008) study—who were amateurs—as “experts.” Moreover, Ericsson and Charness (1994) proposed that someone need only perform two standard deviations above the population mean to be considered an expert. For a theory to be falsifiable, definitions must be used consistently.

[Diagram: Fig. 1. Average proportion of variance in sports performance explained by deliberate practice (light gray region) for (a) all the studies in the meta-analysis (63 effect sizes), (b) the 10 studies that used an objective measure of performance (15 effect sizes), (c) the 10 studies that measured individual practice (no group practice; 14 effect sizes), and (d) the 26 studies that did not use a composite measure that included competition and/or play-like activities (53 effect sizes; Ericsson, 2016, states that 9 of 52 effect sizes were based on this type of composite; it was 10 of 63, see Open Data at osf.io/r5gjw).]

**Other Issues**

Ericsson’s (2016) evaluation of our research is further undermined by omissions and errors. For example, he criticizes us for including nine effect sizes for composites that “were not even pure estimates of different types of practice, [but] ... included hours for play and competition” (p. 352). He fails, though, to mention our report that the deliberate practice–performance relationship was
virtually unchanged with these effect sizes excluded (18% to 17%; Fig. 1d). (The composites consisted of multiple activities interpretable as deliberate practice; hence their inclusion in the main meta-analysis.)

Ericsson (2016) further claims that for one study we “only included the nonsignificant correlation with total sum of all practice (Young, 1998) and disregarded the significant correlations with specialized training activities (Young & Salmela, 2010)” (p. 353). Ericsson is mistaken: We included the effect size for the measure Young (1998) labeled as “deliberate practice”—not the total sum of all practice—and we excluded Young and Salmela’s (2010) study because their sample was a subset of Young’s (1998). We did not, as Ericsson contends, sum up “virtually any type of sports-specific activity” (p. 351).

Ericsson (2016) also criticizes us for excluding an effect size from Law, Côté, and Ericsson (2007). However, as we explained, that effect size fell outside the valid range for correlations (r > 1.0); we also reported supplemental analyses retaining it, and the results are virtually unchanged (Macnamara et al., 2016; Table S2). Finally, Ericsson implies that we did not include an effect size for coach-led instruction from Baker, Côté, and Abernethy (2003). In fact, we did (d = .99 from Baker et al.’s Table 3). Our data file is openly available at osf.io/r5qjw, with all effect sizes and descriptions of measures.

Conclusion
The available evidence indicates that deliberate practice, though undeniably important, does not largely account for individual differences in expertise. Building on Ericsson’s pioneering work, the task now is to develop theories of expertise that include multiple factors.

Declaration of Conflicting Interests
The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References


