Psychological Science OnlineFirst, published on July 1, 2014 as doi:10.1177/0956797614535810

Research Article

Psychological Science
1–11

DOI: 10.1177/0956797614535810
pss.sagepub.com

Deliberate Practice and Performance in Music, Games, Sports, Education, and Professions: A Meta-Analysis

Brooke N. Macnamara[1], David Z. Hambrick[2], and Frederick L. Oswald[3]
[1]Princeton University; [2]Michigan State University; and [3]Rice University

Abstract
More than 20 years ago, researchers proposed that individual differences in performance in such domains as music, sports, and games largely reflect individual differences in amount of *deliberate practice*, which was defined as engagement in structured activities created specifically to improve performance in a domain. This view is a frequent topic of popular-science writing—but is it supported by empirical evidence? To answer this question, we conducted a meta-analysis covering all major domains in which deliberate practice has been investigated. We found that deliberate practice explained 26% of the variance in performance for games, 21% for music, 18% for sports, 4% for education, and less than 1% for professions. We conclude that deliberate practice is important, but not as important as has been argued.

Keywords
deliberate practice, talent development, meta-analysis, human performance, skill acquisition, expertise, open data

Received 8/24/13; Revision accepted 4/24/14

Why do so few people who take up an instrument such as the violin, a sport such as golf, or a game such as chess ever reach an expert level of performance? This question is a topic of a long-running debate in psychology. There are two classical views. One is that experts are "born"—that training is necessary to reach a high level of performance, but innate ability limits the ultimate level of performance a person can achieve. Galton (1869), the founder of behavioral genetics, argued for this position on the basis of his finding that eminence in science, music, art, sports, and other domains tends to run in families. The opposing view is that experts are "made"—that either talent does not exist or its effects on performance are overshadowed by the effect of training. Watson (1930), the founder of behaviorism, captured this view when he stated that "practicing more intensively than others . . . is probably the most reasonable explanation we have today not only for success in any line, but even for genius" (p. 212).

More recently, in the spirit of Watson, Ericsson, Krampe, and Tesch-Römer (1993) proposed their influential *deliberate-practice view* of expert performance. This view holds that expert performance largely reflects accumulated amount of deliberate practice, which Ericsson et al. defined as engagement in structured activities created specifically to improve performance in a domain. In two studies, Ericsson et al. recruited musicians with different levels of accomplishment and asked them to estimate the amount of deliberate practice they had engaged in per week for each year of their musical careers. On average, cumulative amount of deliberate practice was much higher for the most-accomplished groups of musicians than for the less-accomplished groups. For example, at age 20, the average for the "best" violinists was more than 10,000 hr, whereas the averages were about 7,800 hr for the "good" violinists and about 4,600 hr for the least-accomplished group.

Ericsson et al. (1993) concluded that "high levels of deliberate practice are necessary to attain expert level

Corresponding Author:
Brooke N. Macnamara, Department of Psychology, Peretsman-Scully Hall, Princeton University, Princeton, NJ 08544
E-mail: bmacnama@princeton.edu

performance" and added, "Our theoretical framework can *also provide a sufficient account* [emphasis added] of the major facts about the nature and scarcity of exceptional performance. Our account does not depend on scarcity of innate ability (talent) . . ." (p. 392). They continued, "We argue that the differences between expert performers and normal adults reflect a life-long period of deliberate effort to improve performance in a specific domain" (p. 400). Ericsson (2007) reiterated this perspective when he claimed that "the distinctive characteristics of elite performers are adaptations to extended and intense practice activities that selectively activate dormant genes that all healthy children's DNA contain[s]" (p. 4).

The deliberate-practice view has inspired a great deal of interest in expert performance. A Google Scholar search in April 2014 showed that the article by Ericsson et al. (1993) has been cited more than 4,200 times (http://scholar.google.com/scholar?cites=11519303805153777449&as_sdt=20000005&sciodt=0,21&hl=en), and their research has been discussed in a number of popular books, including Gladwell's (2008) *Outliers*, Levitt and Dubner's (2009) *SuperFreakonomics*, and Colvin's (2008) *Talent Is Overrated*. Ericsson et al.'s findings were also the inspiration for what Gladwell termed the "10,000-hour rule"—the idea that it takes 10,000 hr of practice to become an expert.

At the same time, the deliberate-practice view has been sharply criticized in the scientific literature. Gardner (1995) commented that the view requires a "blindness . . . to decades of psychological theorizing" (p. 802), and Sternberg (1996) observed that "deliberate practice may be correlated with success because it is a proxy for ability: We stop doing what we do not do well and feel unrewarded for" (p. 350). Anderson (2000) stated that "Ericsson and Krampe's research does not really establish the case that a great deal of practice is sufficient for great talent" (p. 324), and Marcus (2012) concluded that "it would be a logical error to infer from the importance of practice that talent is somehow irrelevant, as if the two were in mutual opposition" (p. 94).

Furthermore, although deliberate practice is important, growing evidence indicates that it is not as important as Ericsson and colleagues (Ericsson, 2007; Ericsson et al., 1993; Ericsson & Moxley, 2012) have argued. Gobet and Campitelli (2007) found a large amount of variability in total amount of deliberate practice even among master-level chess players—from slightly more than 3,000 hr to more than 23,000 hr. In a recent reanalysis of previous findings, Hambrick et al. (2014) found that deliberate practice accounted for about one third of the reliable variance in performance in chess and music. Thus, in these domains, a large proportion of the variance in performance is explainable by factors other than deliberate practice.

The Current Meta-Analysis

Our meta-analysis is a broad investigation of studies relevant to the deliberate-practice view. It is the first formal meta-analysis of the relationship between deliberate practice and human performance, and we cover all major domains in which this relationship has been studied: music, games, sports, professions, and education.

Our first goal was to estimate the overall correlation between amount of deliberate practice and performance. Ericsson and his colleagues have based their conclusions about the importance of deliberate practice on findings with measures reflecting the accumulated amount (i.e., number of hours) of deliberate practice (e.g., Duffy, Baluch, & Ericsson, 2004; Ericsson et al., 1993; Lehmann & Ericsson, 1996; Tuffiash, Roring, & Ericsson, 2007). Thus, we sought to answer a specific question: How much of the total variance in performance is explained by the accumulated amount of deliberate practice?

Our second goal was to investigate factors that might moderate the relationship between deliberate practice and performance. The first set of factors, which we term *theoretical moderators*, included domain (music, games, sports, professions, or education[1]) and predictability of the task environment (i.e., the degree to which the task environment can change while the performer is planning and executing an action and the range of possible actions). There were three levels of predictability—low, medium, and high. An example of an activity with a low-predictability environment was handling an aviation emergency; an example of an activity with a moderate-predictability environment was the sport of fencing; and an example of an activity with a high-predictability environment was running. We made no prediction about how the strength of the relationship between deliberate practice and performance would vary across domains. However, we did predict that this relationship would generally be more positive for high-predictability activities than for low-predictability activities, on the basis of findings that effects of training on performance are stronger when the task environment is more predictable (e.g., Ackerman, Kanfer, & Goff, 1995; Schneider & Fisk, 1982).

The second set of factors, which we term *methodological moderators*, included (a) the method used to assess deliberate practice—retrospective questionnaire, retrospective interview, or log—and (b) the method used to assess performance—expert rating of performance, standardized objective measure of performance (e.g., chess rating), group membership (e.g., amateur vs. professional), or performance on an objectively scored laboratory task. When a retrospective method is used to assess deliberate practice (questionnaire or interview), participants are asked to recall and estimate their past engagement in deliberate practice. By contrast, when the log method is used, deliberate practice is recorded on an ongoing basis, either by the participant in a diary or by a computer. Given that people do not have perfect memory for the past, particularly the distant past, the log method presumably yields more accurate (valid) estimates of deliberate practice than retrospective methods do. Therefore, we wanted to determine whether the relationship between deliberate practice and performance differed for the log method and for the retrospective methods.

Method

We designed the meta-analysis and report the results in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009).

Inclusion criteria, literature search, and coding

The criteria for including a study in the meta-analysis were as follows:

- A measure of accumulated amount (e.g., number of hours) of one or more activities interpretable as deliberate practice (henceforth, *deliberate practice*) was collected, and the study report referred to at least one publication on deliberate practice by Ericsson and his colleagues.[2]
- A measure of performance reflecting level of skill in the particular domain was collected.
- An effect size reflecting the relationship between accumulated amount of deliberate practice and performance was reported, or information needed to compute this effect size could be obtained from the author(s) of the study.
- The methods and results were in English.
- The participants were human.

We did not exclude studies on the basis of participants' age or skill level.

To identify studies meeting these criteria, we systematically searched for relevant published and unpublished articles in psychology, education, sports science, medicine, and other disciplines through March 24, 2014 (for a flowchart designed according to the PRISMA specifications, see Fig. 1). We also e-mailed authors of articles on deliberate practice and requested information relevant to our meta-analysis that was not accessible (e.g., unpublished data), and we asked that they forward the e-mail to colleagues who might have conducted relevant studies.

Our search and e-mail request yielded 9,331 potentially relevant articles. After examining these articles and discarding irrelevant ones (e.g., literature reviews, commentaries), we identified 88 studies that met all the inclusion criteria. We coded each study and the measures collected in it for reference information, methodological characteristics, and results (the data file is openly available at https://osf.io/rhfsk). These studies included 111 independent samples, with 157 effect sizes and a total sample size of 11,135 participants. For a list of studies included in the meta-analysis, see the Supplemental Method and Results in the Supplemental Material available online. For additional characteristics of the meta-analysis, see Table 1.

The first and second authors coded each effect for moderator variables, and then two individuals with no knowledge of the effect sizes provided separate sets of coding. As indexed by Cohen's kappa for the categorical variables and Spearman's rho for the quantitative variable, interrater agreement among the independent raters and agreement between these individuals' ratings and the authors' ratings were generally high—domain: κs = .99–1.00; predictability of the task environment: ρs = .89–.96; method used to assess deliberate practice: κs = .91–.98; and method used to assess performance: κs = .78–.83. The authors resolved any discrepancies.

Effect sizes

The meta-analysis used the correlation between accumulated amount of deliberate practice and performance as the measure of effect size. For most studies, the authors reported a correlation coefficient;[3] for studies in which the authors reported group-level comparisons (e.g., professional vs. amateur musicians), we converted standardized mean differences (Cohen's *d*s) to biserial correlations (r_bs; Becker, 1986; Hunter & Schmidt, 1990).

Meta-analytic procedure

The meta-analysis involved four steps. The first step was to obtain correlations between time spent in one or more activities interpretable as deliberate practice and performance, along with their sampling error variances. The second step was to search for extreme values. One effect size exceeded 1.0 (r = 1.15); we judged this effect size to be invalid and deleted it. There also were four outliers—effect sizes whose residuals had z scores of 3 or greater (rs = .91, .90, .90, and .84); we Winsorized these values to z scores equaling 2.99 (rs = .83, .83, .84, and .83, respectively). The third step was to estimate overall effects and heterogeneity in the effect sizes using random-effects meta-analysis modeling, and then to test whether some of the heterogeneity was predictable from

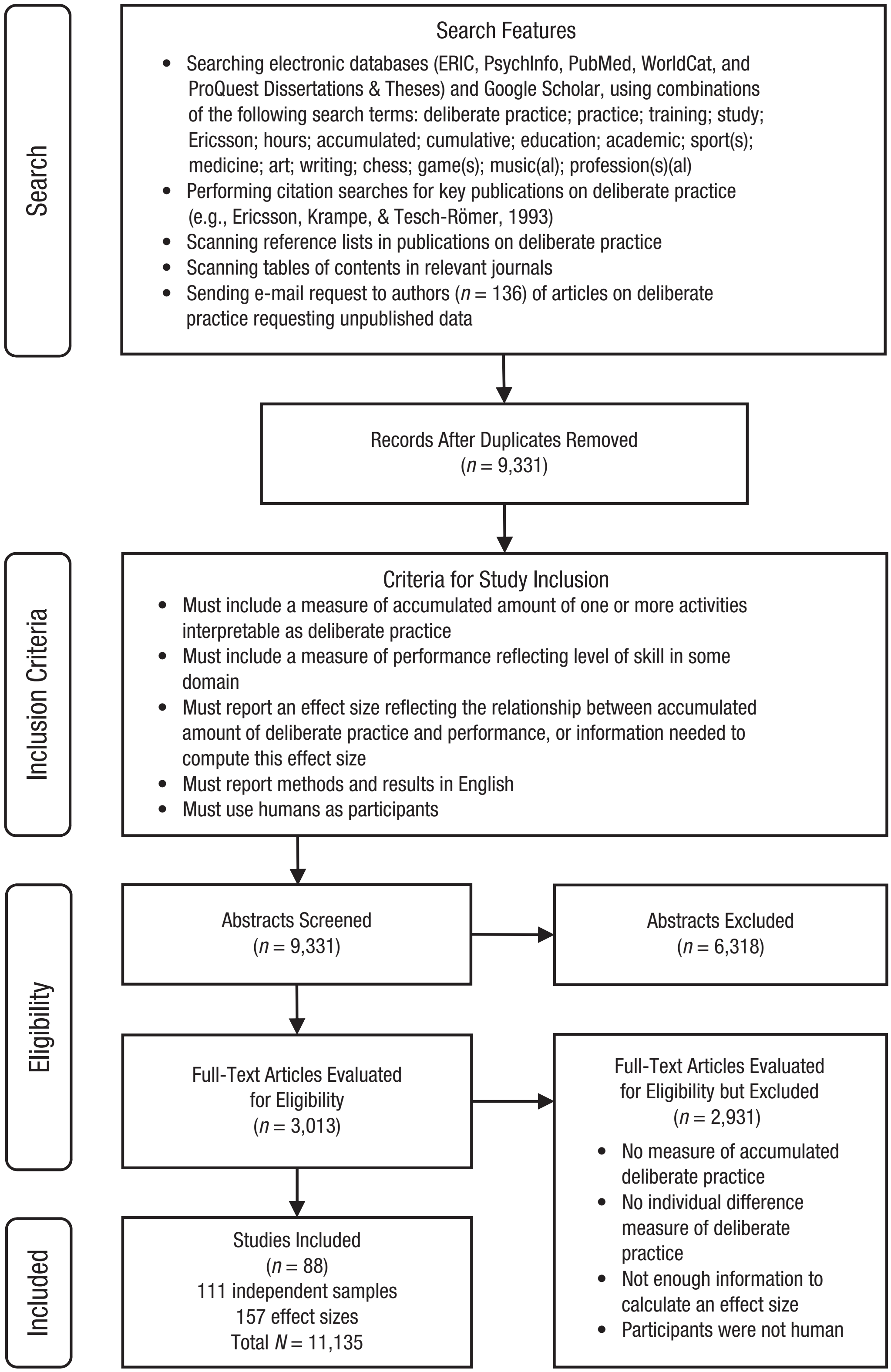


Fig. 1. Flow diagram of the literature search and study coding.

Table 1. Descriptive Characteristics of the Meta-Analysis

Study characteristic	Number of effect sizes	Number of participants
Domain		
Music	28	1,259
Games	11	1,291
Sports	60	2,633
Professions	7	321
Education	51	5,631
Method used to estimate deliberate practice hours		
Interview	36	1,238
Questionnaire	96	8,233
Log	25	1,664
Method used to estimate performance[a]		
Standardized objective measure	73	7,275
Laboratory measure	20	473
Group membership	35	2,266
Expert rating	29	1,307
Publication status		
Published	128	10,155
Unpublished	29	980
Total	157	11,135

[a]For this characteristic, the number of participants does not add up to 11,135 because some samples contributed to multiple types of effects.

moderator variables using mixed-effects meta-analysis modeling. The final step was to perform publication-bias analyses. We used the Comprehensive Meta Analysis (Version 2; Biostat, Englewood, NJ) software package to conduct the meta-analyses and publication-bias analyses. (See also Methodological Details and Screen Shots of Results, Figs. S3–S16, in the Supplemental Method and Results in the Supplemental Material.)

Results

Figure 2 shows that nearly all correlations between deliberate practice and performance were positive: High levels of deliberate practice were associated with high levels of performance. Of the small number of negative correlations (10 of 157), only 2 (< 1.5% of all correlations) were statistically significant ($p < .05$).

The meta-analytic average correlation between deliberate practice and performance was .35, 95% confidence interval (CI) = [.30, .39], which indicates that deliberate practice explained 12% of the variance in performance, 95% CI = [9%, 15%]; thus, 88% of the variance was unexplained. However, as indicated by the I^2 statistic, which specifies the percentage of the between-study variability in effect sizes that is due to heterogeneity rather than random error, there was a high degree of heterogeneity in the effect sizes, $I^2 = 84.90$. We investigated the source of this heterogeneity through the moderator analyses reported next.

Moderator analyses

Theoretical moderators. Domain was a statistically significant moderator, $Q(4) = 49.09$, $p < .001$. Percentage of variance in performance explained by deliberate practice was 26% for games ($\bar{r} = .51$, $p < .001$), 21% for music ($\bar{r} = .46$, $p < .001$), 18% for sports ($\bar{r} = .42$, $p < .001$), 4% for education ($\bar{r} = .21$, $p < .001$), and less than 1% for professions ($\bar{r} = .05$, $p = .62$; see Fig. 3).

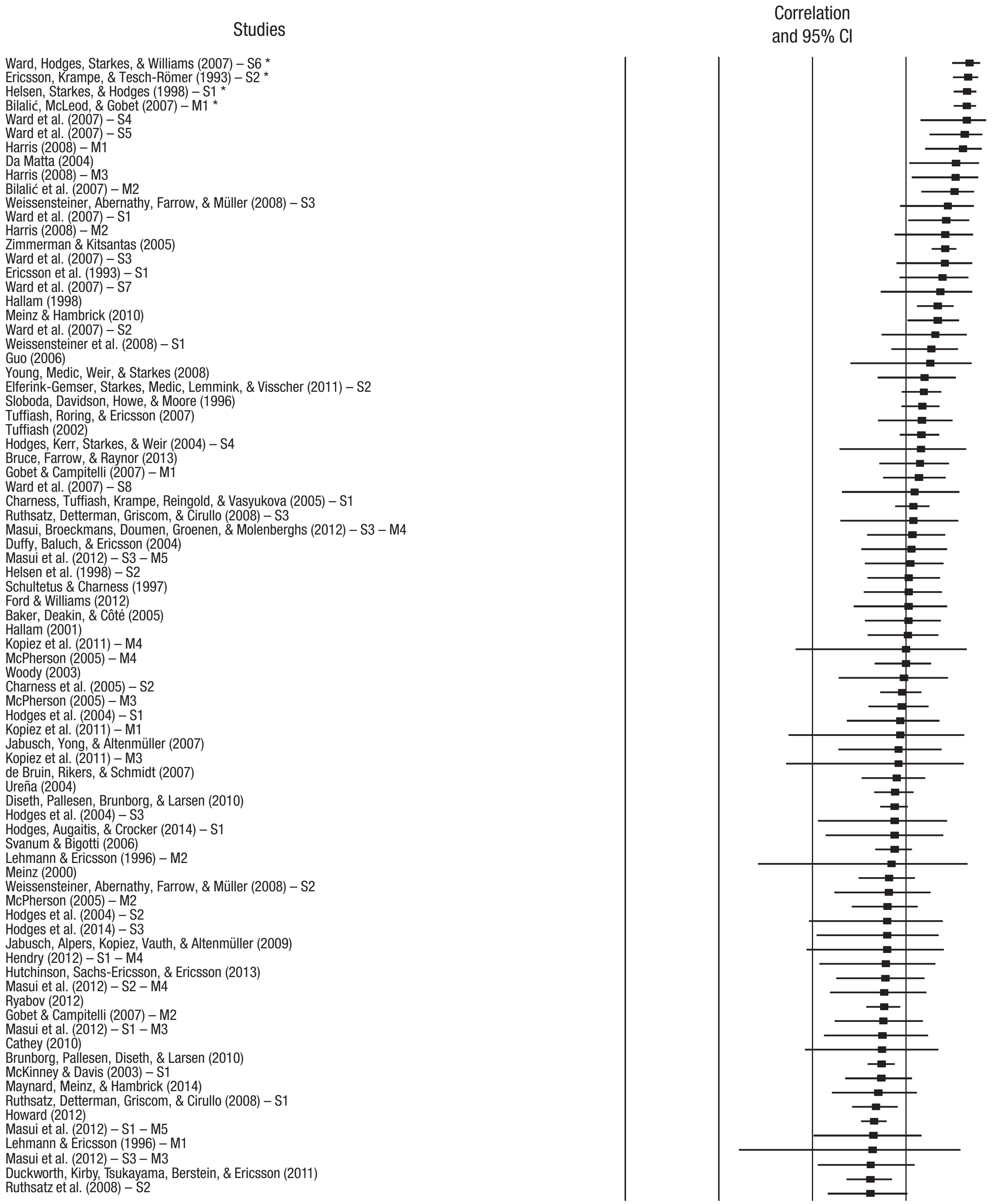
Predictability of the task environment was also a statistically significant moderator, $Q(1) = 20.49$, $b = 0.14$, $T^2 = .05$, $p < .001$. As hypothesized, the percentage of variance in performance explained by deliberate practice was largest (24%) for activities high in predictability ($\bar{r} = .49$), intermediate (12%) for activities moderate in predictability ($\bar{r} = .35$), and smallest (4%) for activities low in predictability ($\bar{r} = .21$; see also Fig. S1 in the Supplemental Method and Results in the Supplemental Material).

Methodological moderators. The method used to assess deliberate practice was a statistically significant moderator, $Q(2) = 16.19$, $p < .001$. The percentage of variance in performance explained by deliberate practice was 20% for studies that used a retrospective interview ($\bar{r} = .45$, $p < .001$), 12% for studies that used a retrospective questionnaire ($\bar{r} = .34$, $p < .001$), and 5% for studies that used a log method ($\bar{r} = .22$, $p < .001$).[4]

The method used to assess performance was also a statistically significant moderator, $Q(3) = 14.41$, $p = .002$. The percentage of variance in performance explained by deliberate practice was 26% for studies that used group membership ($\bar{r} = .51$, $p < .001$), 14% for studies that used laboratory tasks ($\bar{r} = .37$, $p < .001$), 9% for studies that used expert ratings ($\bar{r} = .30$, $p < .001$), and 8% for studies that used standardized objective scoring measures ($\bar{r} = .28$, $p < .001$).

Additional meta-analytic models

We ran three additional models. The first model excluded the 38 effect sizes for team sports, leaving 119 effect sizes (games: 11, music: 28, individual sports: 22, education: 51, professions: 7). We ran this model because interpretation of correlations between deliberate practice and performance in team sports is complicated by the fact that an individual's performance is not independent of the team's performance (Hutchinson, Sachs-Ericsson, & Ericsson, 2013). The overall percentage of variance explained by deliberate practice was 11% in this model


Studies
Correlation
and 95% CI
Ward, Hodges, Starkes, & Williams (2007) – S6 *
Ericsson, Krampe, & Tesch-Römer (1993) – S2 *
Helsen, Starkes, & Hodges (1998) – S1 *
Bilalić, McLeod, & Gobet (2007) – M1 *
Ward et al. (2007) – S4
Ward et al. (2007) – S5
Harris (2008) – M1
Da Matta (2004)
Harris (2008) – M3
Bilalić et al. (2007) – M2
Weissensteiner, Abernathy, Farrow, & Müller (2008) – S3
Ward et al. (2007) – S1
Harris (2008) – M2
Zimmerman & Kitsantas (2005)
Ward et al. (2007) – S3
Ericsson et al. (1993) – S1
Ward et al. (2007) – S7
Hallam (1998)
Meinz & Hambrick (2010)
Ward et al. (2007) – S2
Weissensteiner et al. (2008) – S1
Guo (2006)
Young, Medic, Weir, & Starkes (2008)
Elferink-Gemser, Starkes, Medic, Lemmink, & Visscher (2011) – S2
Sloboda, Davidson, Howe, & Moore (1996)
Tuffiash, Roring, & Ericsson (2007)
Tuffiash (2002)
Hodges, Kerr, Starkes, & Weir (2004) – S4
Bruce, Farrow, & Raynor (2013)
Gobet & Campitelli (2007) – M1
Ward et al. (2007) – S8
Charness, Tuffiash, Krampe, Reingold, & Vasyukova (2005) – S1
Ruthsatz, Detterman, Griscom, & Cirullo (2008) – S3
Masui, Broeckmans, Doumen, Groenen, & Molenberghs (2012) – S3 – M4
Duffy, Baluch, & Ericsson (2004)
Masui et al. (2012) – S3 – M5
Helsen et al. (1998) – S2
Schultetus & Charness (1997)
Ford & Williams (2012)
Baker, Deakin, & Côté (2005)
Hallam (2001)
Kopiez et al. (2011) – M4
McPherson (2005) – M4
Woody (2003)
Charness et al. (2005) – S2
McPherson (2005) – M3
Hodges et al. (2004) – S1
Kopiez et al. (2011) – M1
Jabusch, Yong, & Altenmüller (2007)
Kopiez et al. (2011) – M3
de Bruin, Rikers, & Schmidt (2007)
Ureña (2004)
Diseth, Pallesen, Brunborg, & Larsen (2010)
Hodges et al. (2004) – S3
Hodges, Augaitis, & Crocker (2014) – S1
Svanum & Bigotti (2006)
Lehmann & Ericsson (1996) – M2
Meinz (2000)
Weissensteiner, Abernathy, Farrow, & Müller (2008) – S2
McPherson (2005) – M2
Hodges et al. (2004) – S2
Hodges et al. (2014) – S3
Jabusch, Alpers, Kopiez, Vauth, & Altenmüller (2009)
Hendry (2012) – S1 – M4
Hutchinson, Sachs-Ericsson, & Ericsson (2013)
Masui et al. (2012) – S2 – M4
Ryabov (2012)
Gobet & Campitelli (2007) – M2
Masui et al. (2012) – S1 – M3
Cathey (2010)
Brunborg, Pallesen, Diseth, & Larsen (2010)
McKinney & Davis (2003) – S1
Maynard, Meinz, & Hambrick (2014)
Ruthsatz, Detterman, Griscom, & Cirullo (2008) – S1
Howard (2012)
Masui et al. (2012) – S1 – M5
Lehmann & Ericsson (1996) – M1
Masui et al. (2012) – S3 – M3
Duckworth, Kirby, Tsukayama, Berstein, & Ericsson (2011)
Ruthsatz et al. (2008) – S2

(continued)

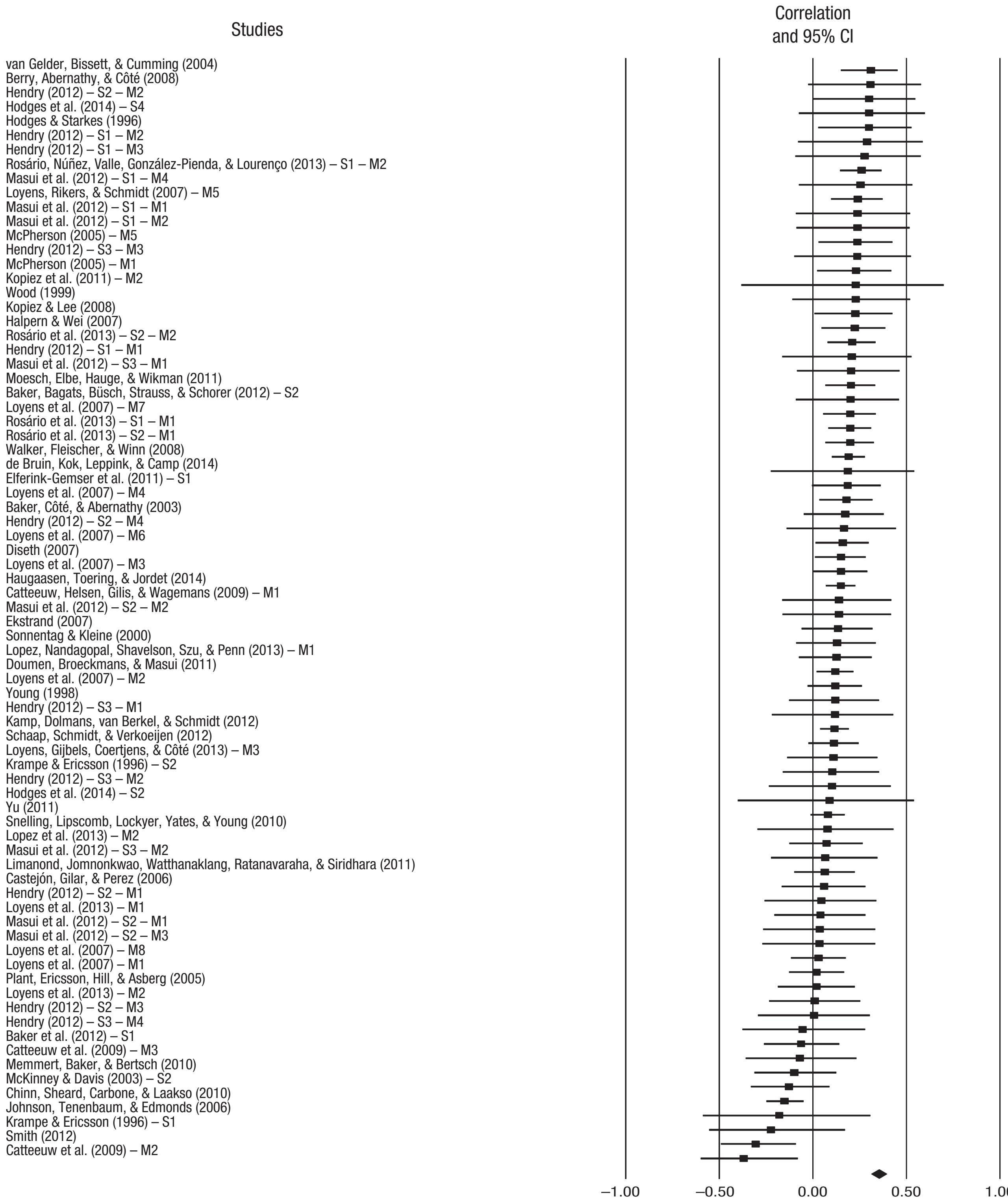


Fig. 2. Correlations between deliberate practice and performance. Correlations (squares) and 95% confidence intervals (CIs; lines) are displayed for all effects entered into the meta-analysis. The diamond on the bottom row represents the meta-analytically weighted mean correlation. Multiple measures were adjusted for dependency (see also Methodological Details in the Supplemental Method and Results in the Supplemental Material). Asterisks identify adjusted (Winsorized) outliers. For studies with multiple independent samples, the result for each sample (S1, S2, etc.) is reported separately. Similarly, for studies with multiple performance measures, the result for each measure (M1, M2, etc.) is reported separately.

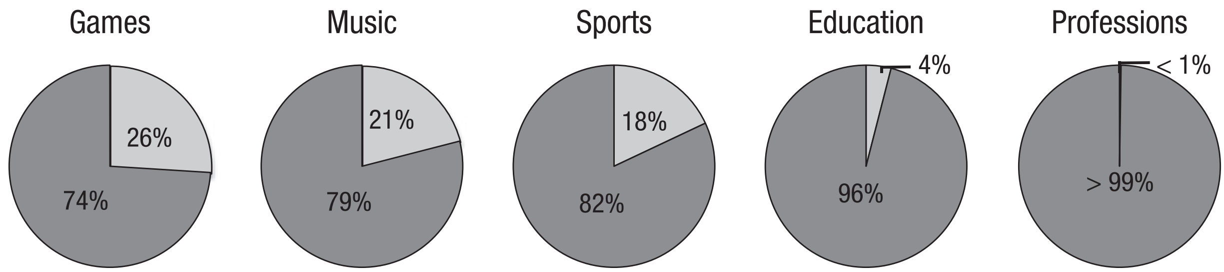


Fig. 3. Percentage of variance in performance explained (light gray) and not explained (dark gray) by deliberate practice within each domain studied. Percentage of variance explained is equal to $r^2 \times 100$.

(games: 26%, music: 21%, sports: 19%, and education: 4%, all *p*s < .001; professions: < 1%, *p* = .62).

The second model included only the 59 effect sizes for solitary deliberate practice (games: 6; music: 9; sports: 14; education: 30; professions: 0). We tested this model to address the question of whether deliberate practice must be performed in isolation to be maximally effective (Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005; Ericsson et al., 1993). The overall percentage of variance explained by deliberate practice was 11% in this model (games: 23%; music: 23%; sports: 22%; and education: 3%; all *p*s < .001), which indicates that solitary deliberate practice is not a stronger predictor of performance than deliberate practice with other people.

The third model included only the 53 effect sizes for solitary deliberate practice available after excluding effect sizes for team sports (games: 6; music: 9; individual sports: 8; education: 30). The overall percentage of variance explained by deliberate practice was 10% in this model (games: 23%; music: 23%; sports: 28%; and education: 3%; all *p*s < .001).

Thus, results of the additional analyses were similar and consistent with the overall analysis, indicating that deliberate practice explained a considerable amount of the variance in performance, but a large amount of the variance remains unexplained.

Publication-bias analyses

We conducted publication-bias analyses to investigate whether null or weak results have been systematically suppressed from publication in the deliberate-practice literature and whether there were effect sizes missing from our meta-analysis because of publication bias. We first inspected a funnel plot depicting the relationship between standard error and effect size; it was approximately symmetrical, suggesting that smaller-sample studies with weak effect sizes were not missing from our meta-analysis (see Fig. S2 in Additional Publication-Bias Analyses in the Supplemental Method and Results in the Supplemental Material). A trim-and-fill analysis (Duval & Tweedie, 2000a, 2000b) confirmed this, indicating that no effects were missing because of publication bias.

General Discussion

More than 20 years ago, Ericsson et al. (1993) argued that "individual differences in ultimate performance can largely be accounted for by differential amounts of past and current levels of practice" (p. 392). Ericsson and Moxley (2012) reiterated this claim, stating that "the concept of deliberate practice can account for the large individual differences between experts and novices" (p. 145). The results of this meta-analysis do not support these strong claims. Regardless of domain, a large amount of variance in performance is not explained by deliberate practice and is potentially explainable by other factors. We conclude that amount of deliberate practice—although unquestionably important as a predictor of individual differences in performance from both a statistical and a practical perspective—is not as important as Ericsson and his colleagues have argued.

Moderator analyses revealed that the strength of the relationship between deliberate practice and performance varied by domain. In terms of percentage of variance in performance explained, the effect of deliberate practice was strong for games (26%), music (21%), and sports (18%), and much weaker for education (4%) and professions (< 1% and not statistically significant). Why were the effect sizes for education and professions so much smaller? One possibility is that deliberate practice is less well defined in these domains. It could also be that in some of the studies, participants differed in amount of prestudy expertise (e.g., amount of domain knowledge before taking an academic course or accepting a job) and thus in the amount of deliberate practice they needed to achieve a given level of performance.

Moderator analyses further revealed that the effect of deliberate practice on performance tended to be larger for activities that are highly predictable (e.g., running) than for activities that are less predictable (e.g., handling an aviation emergency), as we hypothesized. Furthermore, the effect of deliberate practice on performance was stronger for studies that used retrospective methods to elicit estimates of deliberate practice than for those that used a log method. In fact, for studies using the log method, which presumably yields more valid estimates than retrospective methods do, deliberate practice accounted for only 5% of the variance in performance. This finding suggests that the use of what Ericsson (2014) termed a "high-fidelity" (p. 13) approach to assessing deliberate practice (e.g., video monitoring) might reveal that the relationship between deliberate practice and performance is weaker than the results of this meta-analysis indicate. Finally, the relationship between deliberate practice and performance was weaker for studies that used a standardized objective measure of performance (e.g., chess rating) than for studies that used group membership as the measure of performance.

We did not correct individual effect sizes for the attenuating effect of measurement error (i.e., measurement unreliability), because very few studies in the meta-analysis reported a reliability estimate for both deliberate practice and performance. However, measures of both deliberate practice and performance are typically found to have acceptable or better reliability (≥ .70). For example, Tuffiash et al. (2007) stated that test-retest reliabilities for self-report practice estimates in sports and music are typically at or above .80, and Hambrick et al. (2014) found reliability of .91 for chess ratings. Furthermore, the percentage of variance in performance explained by deliberate practice is smaller than the percentage of variance not explained by deliberate practice[5] across a wide range of reliability assumptions (see Table S1 in the Supplemental Method and Results in the Supplemental Material). For example, if it is assumed that reliability of both deliberate practice and performance is .80, the mean overall correlation between deliberate practice and performance is .43 after correction for unreliability. This correlation indicates that deliberate practice accounts for 19% of the reliable variance and that 81% of the reliable variance is potentially explainable by other factors; corresponding percentages of variance explained are 41% for games, 33% for music, 28% for sports, 7% for education, and less than 1% for professions.

What explains the variance in performance that deliberate practice does not explain? There are probably many factors. One may be the age at which a person starts serious involvement in a domain. Ericsson et al. (1993) argued that any performance advantage associated with starting age simply reflects the fact that a person who starts at a young age has more time to accumulate deliberate practice than a person who starts at a later age. However, Gobet and Campitelli (2007) and Howard (2012) found that starting age negatively predicted chess rating even after statistically controlling for deliberate practice. This evidence suggests that there may be an optimal developmental period for acquiring complex skills, as there seems to be for acquiring language (Lenneberg, 1967).

Research suggests that general intelligence and more specific abilities may also explain some of the variance in performance that deliberate practice does not. General intelligence (Hunt, 2011; Jensen, 1998)—which is highly stable and substantially heritable (Plomin, DeFries, McClearn, & McGuffin, 2008)—positively predicts performance in a wide range of domains, including music (Shuter, 1968), chess (Grabner, Stern, & Neubauer, 2007), academics (e.g., Brody, 1997; Laidra, Pullmann, & Allik, 2007), and virtually any occupation (Schmidt & Hunter, 1998, 2004). Working memory capacity—the ability to maintain information in the focus of attention (Engle, 2002)—is an example of a specific ability that may predict performance differences. Meinz and Hambrick (2010) found that working memory capacity positively predicted pianists' performance in a sight-reading task, above and beyond deliberate practice. There was no significant interaction between deliberate practice and working memory capacity, which indicates that working memory capacity was as important a predictor of performance for beginning pianists as it was for pianists who had engaged in thousands of hours of deliberate practice.

Conclusion

Ericsson and his colleagues' (1993) deliberate-practice view has generated a great deal of interest in expert performance, but their claim that individual differences in performance are largely accounted for by individual differences in amount of deliberate practice is not supported by the available empirical evidence. An important goal for future research on expert performance is to draw on existing theories of individual differences (e.g., Ackerman, 1987; Gagné, 2009; Schmidt, 2014; Simonton, 2014) to identify basic abilities and other individual difference factors that explain variance in performance and to estimate their importance as predictor variables relative to deliberate practice. Another important goal is to continue to investigate how and when task and situational factors such as task predictability moderate the impact of deliberate practice and other individual difference factors on performance. Research aimed at addressing these goals will shed new light on the underpinnings of expert performance.

Author Contributions

B. N. Macnamara and D. Z. Hambrick developed the study concept. All authors contributed to the study design. B. N. Macnamara performed effect-size data collection with input from D. Z. Hambrick. B. N. Macnamara performed the data analyses with input and guidance from F. L. Oswald. D. Z.

Hambrick drafted the introduction, discussion, and conclusion sections of the manuscript. B. N. Macnamara drafted the method and results sections. All three authors provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Neil Charness, Donald Chinn, Sarah Doumen, Joakim Ekstrand, Marije Elferink-Gemser, Mathias Haugaasen, Nicola Hodges, Rachelle Kamp, Thirayoot Limanond, Enrique Lopez, Gary McPherson, Joanne Ruthsatz, Igor Ryabov, Lydia Schaap, Steve Schultetus, Judy Sheard, Soren Svanum, and Marjorie Torenbeek for providing unpublished data, additional data and results associated with publications, and clarification of methods or results.

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.

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

Open Practices

All data have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/rhfsk. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/supplemental-data. This article has received the badge for Open Data. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/view/ and http://pss.sagepub.com/content/25/1/3.full.

Notes

1. Studies of professional athletes were included in the sports category, and studies of professional musicians were included in the music category; the professions category included professions not captured by the other domains: computer programming, military aircraft piloting, soccer refereeing, and insurance selling. Studies included in the education category were primarily studies of university students in which the achievement outcome was a course grade or semester grade point average. See the performance measure descriptions in Column N of the Open Data file available at https://osf.io/rhfsk. We classified ballet as a sport because it is a highly physical activity and has similarities to sports such as gymnastics and figure skating.

2. For studies in which the total amount of time that participants had to accumulate deliberate practice was a constant (e.g., a college semester), we were able to use weekly amount of deliberate practice as a measure of accumulated amount of deliberate practice, given that this variable and accumulated amount of deliberate practice would necessarily have the same correlation with performance. The focus of this meta-analysis was on the relationship between individual differences in accumulated deliberate practice and performance. We did not include studies that experimentally manipulated training and then compared trained and untrained individuals.

3. We reversed the sign of the correlation when appropriate before analyzing the data. For instance, negative correlations between deliberate practice and race times in sports indicate that more deliberate practice is associated with lower (faster) race times (i.e., more deliberate practice is associated with better performance).

4. Whether or not the researchers performed a transformation (e.g., log) on the deliberate-practice variable before performing analyses was not a statistically significant moderator of the relationship between deliberate practice and performance, $Q(1) = 1.77$, $p = .18$.

5. The standard formula for correcting a correlation for measurement unreliability is $\hat{r} = r_{xy}/(r_{xx}r_{yy})^{1/2}$, where r_{xx} and r_{yy} are reliability coefficients for x and y, respectively (Schmidt & Hunter, 1996).

References

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

Ackerman, P. L., Kanfer, R., & Goff, M. (1995). Cognitive and noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied*, *1*, 270–304. doi:10.1037/1076-898X.1.4.270

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

Becker, G. (1986). Correcting the point-biserial correlation for attenuation owing to unequal sample size. *The Journal of Experimental Education*, *55*, 5–8.

Brody, N. (1997). Intelligence, schooling, and society. *American Psychologist*, *52*, 1046–1050. doi:10.1037/0003-066X.52.10.1046

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. doi:10.1002/acp.1106

Colvin, G. (2008). *Talent is overrated: What really separates world-class performers from everybody else*. New York, NY: Penguin Group.

Duffy, L. J., Baluch, B., & Ericsson, K. A. (2004). Dart performance as a function of facets of practice amongst professional and amateur men and women players. *International Journal of Sport Psychology*, *35*, 232–245.

Duval, S. J., & Tweedie, R. L. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*, 89–98. doi:10.1080/01621459.2000.10473905

Duval, S. J., & Tweedie, R. L. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463. doi:10.1111/j.0006-341X.2000.00455.x

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

Ericsson, K. A. (2007). Deliberate practice and the modifiability of the 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. (2014). Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms. *Intelligence*, *45*, 81–103. doi: 0.1016/j.intell.2013.12.001

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

Ericsson, K. A., & Moxley, J. H. (2012). The expert performance approach and deliberate practice: Some potential implications for studying creative performance in organizations. In M. D. Mumford (Ed.), *The handbook of organizational creativity* (pp. 141–167). London, England: Academic Press.

Gagné, F. (2009). Talent development as seen through the differentiated model of giftedness and talent. In T. Balchin, B. Hymer, & D. Matthews (Eds.), *The Routledge international companion to gifted education* (pp. 32–41). Abingdon, Oxfordshire, England: Routledge.

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

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

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

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

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

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34–45. doi:10.1016/j.intell.2013.04.001

Howard, R. W. (2012). Longitudinal effects of different types of practice on the development of chess expertise. *Applied Cognitive Psychology*, *26*, 359–369. doi:10.1002/acp.1834

Hunt, E. (2011). *Human intelligence*. Cambridge, England: Cambridge University Press.

Hunter, J. E., & Schmidt, F. L. (1990). Dichotomization of continuous variables: The implications for meta-analysis. *Journal of Applied Psychology*, *75*, 334–339. doi:10.1037/0021-9010.75.3.334

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. doi:10.1080/13598139.2013.780966

Jensen, A. R. (1998). *The* g *factor: The science of mental ability*. Westport, CT: Praeger.

Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, *42*, 441–451. doi:10.1016/j.paid.2006.08.001

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

Lenneberg, E. H. (1967). *Biological foundations of language*. Oxford, England: Wiley.

Levitt, S. D., & Dubner, S. J. (2009). *SuperFreakonomics: Global cooling, patriotic prostitutes, and why suicide bombers should buy life insurance*. New York, NY: HarperCollins.

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

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. doi:10.1177/0956797610373933

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. *PLoS Med* *6*: Article e1000097. Available at http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.1000097

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

Schmidt, F. L. (2014). A general theoretical integrative model of individual differences in interests, abilities, personality traits, and academic and occupational achievement: A commentary on four recent articles. *Perspectives on Psychological Science*, *9*, 211–218. doi:10.1177/1745691613518074

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. doi:10.1037/0022-3514.86.1.162

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, *1*, 199–223. doi:10.1037/1082-989X.1.2.199

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262–274. doi:10.1037/0033-2909.124.2.262

Schneider, W., & Fisk, A. D. (1982). Degree of consistent training: Improvements in search performance and automatic process development. *Perception & Psychophysics*, *31*, 160–168. doi:10.3758/BF03206216

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

Simonton, D. K. (2014). Creative performance, expertise acquisition, individual differences, and developmental antecedents: An integrative research agenda. *Intelligence*, *45*, 66–73. doi:10.1016/j.intell.2013.04.007

Sternberg, R. J. (1996). The costs of expertise. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 347–354). Mahwah, NJ: Erlbaum.

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

Watson, J. B. (1930). *Behaviorism*. New York, NY: W. W. Norton.

Psychological Science
2018, Vol. 29(7) 1202–1204

DOI: 10.1177/0956797618769891
www.psychologicalscience.org/PS

Corrigendum: Deliberate Practice and Performance in Music, Games, Sports, Education, and Professions: A Meta-Analysis

Original article: Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science, 25*, 1608–1618. doi:10.1177/0956797614535810

In our original analysis, we adjusted for dependent performance measures using a method based on Cheung and Chan's (2004, 2008) method. Cheung and Chan's method adjusts the sample size to be between the sample N and the cumulative sample N and applies this to the *average* of the dependent effect sizes. Their adjustment formula is as follows: adjusted $N = ((N - 1)/C) + 1$, where C accounts for the correlation between dependent effect sizes, in addition to the overall average effect size and the number of dependent effect sizes per sample. We inadvertently used the formula as follows: adjusted $N = (N - 1)/(C + 1)$ and then applied this formula to each *individual* effect size (rather than to the average).

It happens that our lower N computed for each individual effect generally offset the higher N for the cumulative effect under the Cheung and Chan approach. The formula and application we used produces values similar to those produced by robust variance estimation (another option for adjusting dependent samples). There was no practical effect on the results when we reanalyzed our findings using Cheung and Chan's approach, and the changes had no impact whatsoever on the substance of our findings and conclusions. Table 1 provides a complete list of the originally reported results that are being changed with this Corrigendum. Questions can be directed to Brooke N. Macnamara (bnm24@case.edu).

Table 1. Results Being Corrected

Result	Originally reported results	Corrected results using Cheung and Chan (2004, 2008) adjustment
	Main model (p. 1612)	
Meta-analytic average correlation between deliberate practice and performance	.35, 95% CI = [.30, .39]	.38, 95% CI = [.33, .42]
Overall variance in performance explained by deliberate practice	12%, 95% CI = [9%, 15%] (88% of variance unexplained)	14%, 95% CI = [11%, 18%] (86% of variance unexplained)
I^2	84.90	88.54
Domain as a moderator	$Q(4) = 49.09$, $p < .001$	$Q(4) = 36.61$, $p < .001$
Variance in performance explained by deliberate practice for games	26% ($\bar{r} = .51$, $p < .001$)	24% ($\bar{r} = .49$, $p < .001$)
Variance in performance explained by deliberate practice for music	21% ($\bar{r} = .46$, $p < .001$)	23% ($\bar{r} = .48$, $p < .001$)
Variance in performance explained by deliberate practice for sports	18% ($\bar{r} = .42$, $p < .001$)	20% ($\bar{r} = .45$, $p < .001$)
Variance in performance explained by deliberate practice for education	4% ($\bar{r} = .21$, $p < .001$)	5% ($\bar{r} = .22$, $p < .001$)
Variance in performance explained by deliberate practice for professions	< 1% ($\bar{r} = .05$, $p = .62$)	1% ($\bar{r} = .09$, $p = .377$)

(*continued*)

Table 1. (*continued*)

Result	Originally reported results	Corrected results using Cheung and Chan (2004, 2008) adjustment
Predictability of the task environment as a moderator	$Q(1) = 20.49$, $b = 0.14$, $T^2 = .05$, $p < .001$	$Q(1) = 11.32$, $b = 0.12$, $T^2 = .05$, $p < .001$
Variance in performance explained by deliberate practice for activities high in predictability	24% ($\bar{r} = .49$)	23% ($\bar{r} = .48$)
Variance in performance explained by deliberate practice for activities moderate in predictability	12% ($\bar{r} = .35$)	14% ($\bar{r} = .37$)
Variance in performance explained by deliberate practice for activities low in predictability	4% ($\bar{r} = .21$)	6% ($\bar{r} = .25$)
Method used to assess deliberate practice as a moderator	$Q(2) = 16.19$, $p < .001$	$Q(2) = 18.18$, $p < .001$
Variance in performance explained by deliberate practice for studies that used a questionnaire	12% ($\bar{r} = .34$, $p < .001$)	15% ($\bar{r} = .38$, $p < .001$)
Variance in performance explained by deliberate practice for studies that used a log method	5% ($\bar{r} = .22$, $p < .001$)	4% ($\bar{r} = .21$, $p < .001$)
Method used to assess performance as a moderator	$Q(3) = 14.41$, $p = .002$	$Q(3) = 9.75$, $p = .021$
Variance in performance explained by deliberate practice for studies that used laboratory tasks	14% ($\bar{r} = .37$, $p < .001$)	12% ($\bar{r} = .35$, $p = .012$)
Variance in performance explained by deliberate practice for studies that used expert ratings	9% ($\bar{r} = .30$, $p < .001$)	11% ($\bar{r} = .34$, $p < .001$)
Variance in performance explained by deliberate practice for studies that used standardized objective scores	8% ($\bar{r} = .28$, $p < .001$)	10% ($\bar{r} = .32$, $p < .001$)
Additional model excluding team sports (pp. 1612, 1615)		
Overall variance in performance explained by deliberate practice	11%	12%
Variance in performance explained by deliberate practice for games	26%, $p < .001$	24%, $p < .001$
Variance in performance explained by deliberate practice for music	21%, $p < .001$	23%, $p < .001$
Variance in performance explained by deliberate practice for sports	19%, $p < .001$	16%, $p < .001$
Variance in performance explained by deliberate practice for education	4%, $p < .001$	5%, $p < .001$
Variance in performance explained by deliberate practice for professions	< 1%, $p = .62$	1%, $p = .377$
Additional model with only solitary deliberate practice (p. 1615)		
Overall variance in performance explained by deliberate practice	11%	14%
Variance in performance explained by deliberate practice for games	23%, $p < .001$	22%, $p < .001$
Variance in performance explained by deliberate practice for music	23%, $p < .001$	25%, $p < .001$
Variance in performance explained by deliberate practice for sports	22%, $p < .001$	18%, $p < .001$
Additional model with only solitary practice and excluding team sports (p. 1615)		
Overall variance in performance explained by deliberate practice	10%	14%
Variance in performance explained by deliberate practice for games	23%, $p < .001$	22%, $p < .001$
Variance in performance explained by deliberate practice for music	23%, $p < .001$	25%, $p < .001$
Variance in performance explained by deliberate practice for sports	28%, $p < .001$	21%, $p < .001$

Note: CI = confidence interval.

References

Cheung, S. F., & Chan, D. K.-S. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology*, *89*, 780–791.

Cheung, S. F., & Chan, D. K.-S. (2008). Dependent correlations in meta-analysis: The case of heterogeneous dependence. *Educational and Psychological Measurement*, *68*, 760–777.

DOI: 10.1177/0956797614535810

Supplemental Online Materials – Reviewed (SOM-R)

Here you will find the list of references for the articles included in the meta-analysis, methodological details of the meta-analysis in an annotated format, additional publication bias analyses, and screen shots of our results from the meta-analysis software package we used to conduct the analyses: Comprehensive Meta-Analysis (Borenstein, Hedges, Higgins, & Rothstein, 2005). For details associated with each effect size, see the Open Data file at https://osf.io/rhfsk.

Studies Included in the Meta-Analysis

Baker, J., Bagats, S., Büsch, D., Strauss, B., & Schorer, J. (2012). Training differences and selection in a talent identification system. *Talent Development & Excellence, 4*, 23–32.

Baker, J., Côté, J., & Abernathy, B. (2003). Learning from the experts: Practice activities of expert decision-makers in sport. *Research Quarterly for Exercise and Sport, 74*, 342–347. doi:10.1080/02701367.2003.10609101

Baker, J., Deakin, J., & Côté, J. (2005). On the utility of deliberate practice: Predicting performance in ultra-endurance triathletes from training indices. *International Journal of Sport Psychology, 36*, 225–240.

Berry, J., Abernathy, B., & Cote, J. (2008). The contribution of structured activity and deliberate play to the development of expert perceptual and decision-making skill. *Journal of Sport & Exercise Psychology, 30*, 685–708.

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

Bruce, L., Farrow, D., & Raynor, A. (2013). Performance milestones in the development of expertise: Are they critical? *Journal of Applied Sport Psychology, 25*, 281–297. doi:10.1080/10413200.2012.725704

Brunborg, G. S., Pallesen, S., Diseth, A., & Larsen, S. (2010). Preoccupation with failure affects number of study hours—Not academic achievement. *Scandinavian Journal of Educational Research, 54*, 125–132. doi:10.1080/00313831003637899

Castejón, J. L., Gilar, R., & Pérez, A. M. (2006). Complex learning: The role of knowledge, intelligence, motivation and learning strategies. *Psicothema, 18*, 679–685.

Cathey, R. M. (2010). *Retrospective practice histories of expert and novice baseball pitchers* (Doctoral dissertation). University of South Carolina, Columbia. Available from ProQuest Dissertations and Theses database. (UMI No. 3413286)

Catteeuw, P., Helsen, W., Gilis, B., & Wagemans, J. (2009). Decision-making skills, role specificity, and deliberate practice in association football refereeing. *Journal of Sports Sciences, 27*, 1125–1136. doi:10.1080/02640410903079179

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. doi:10.1002/acp.1106

Chinn, D., Sheard, J., Carbone, A., & Laakso, M.-J. (2010). Study habits of CS 1 students: What do they do outside the classroom? In T. Clear & J. Hamer (Eds.), *Proceedings of the Twelfth Australasian Computing Education Conference (AAQ 2010)* (CRPIT, Vol. 103, pp. 53–62). Brisbane, Australia: Australian Computer Society.

Da Matta, G. B. (2004). *The influence of deliberate practice and social support systems on the development of expert and intermediate women volleyball players in Brazil* (Doctoral dissertation). University of South Carolina, Columbia. Available from ProQuest Dissertations and Theses database. (UMI No. 3142809)

de Bruin, A. B. H., Kok, E. M., Leppink, J. M., & Camp, G. (2014). Practice, intelligence, and enjoyment in novice chess players: A prospective study at the earliest stage of a chess career. Intelligence, 45, 18–25. doi:10.1016/j.intell.2013.07.004

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

Diseth, Å. (2007). Approaches to learning, course experience and examination grade among undergraduate psychology students: Testing of mediator effects and construct validity. *Studies in Higher Education, 32*, 373–388. doi:10.1080/03075070701346949

Diseth, Å., Pallesen, S., Brunborg, G. S., & Larsen, S. (2010). Academic achievement among first semester undergraduate psychology students: The role of course experience, effort, motives and learning strategies. *Journal of Higher Education, 59*, 335–352. doi:10.1007/s10734-009-9251-8

Doumen, S., Broeckmans, J., & Masui, C. (2011, June 29–July 1). *Study time and academic performance: A conditional relation?* Paper presented at the Annual Conference of the Education, Learning, Styles, Individual Differences Network, Antwerp, Belgium.

Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate practice spells success: Why grittier competitors triumph at the national spelling bee. *Social Psychological & Personality Science, 2*, 174–181. doi:10.1177/1948550610385872

Duffy, L. J., Baluch, B., & Ericsson, K. A. (2004). Dart performance as a function of facets of practice amongst professional and amateur men and women players. *International Journal of Sport Psychology, 35*, 232–245.

Ekstrand, J. (2007). *Measuring and explaining economics students learning efficiency*. Unpublished document. Retrieved from http://www.nhh.no/Files/Filer/institutter/for/conferences/nff/papers/ekstrand.pdf

Elferink-Gemser, M. T., Starkes, J. L., Medic, N., Lemmink, K. A. P. M., & Visscher, C. (2011). What discriminates elite and sub-elite youth field hockey players? *Annals of Research in Sport and Physical Activity, 2*, 50–68. doi:10.1080/02640410410001729991

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

Ford, P. R., & Williams, A. M. (2012). The developmental activities engaged in by elite youth soccer players who progressed to professional status compared to those who did not. *Psychology of Sport Exercise, 13*, 1–4. doi:10.1016/j.psychsport.2011.09.004

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

Guo, A. (2006). Competition preparation and deliberate practice: A study of the 2005 National Spelling Bee finalists. *Dissertation Abstracts International, 67*(09), 3297. (Pub. No. 3235947)

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

Hallam, S. (2001). The development of metacognition in musicians: Implications for education. *British Journal of Music Education, 18*, 27–39. doi:10.1017/S0265051701000122

Halpern, D. F., & Wai, J. (2007). The world of competitive scrabble: Novice and expert differences in visuospatial and verbal abilities. *Journal of Experimental Psychology: Applied, 13*, 79–94. doi:10.1037/1076-898X.13.2.79

Harris, K. R. (2008). *Deliberate practice, mental representations, and skilled performance in bowling* (Doctoral dissertation). Florida State University. Available from Electronic Theses, Treatises and Dissertations, Diginole Commons. (Paper No. 4245)

Haugaasen, M., Toering, T., & Jordet, G. (2014). From childhood to senior professional football: A multi-level approach to elite youth football player's engagement in football specific activities. *Psychology of Sport and Exercise, 15,* 336–344. doi:10.1016/j.psychsport.2014.02.007

Helsen, W. F., Starkes, J. L., & Hodges, N. J. (1998). Team sports and the theory of deliberate practice. *Journal of Sport & Exercise Psychology, 20*, 12–34.

Hendry, D. T. (2012). *The role of developmental activities on self determined motivation, passion and skill in youth soccer players* (Master's thesis). The University of British Columbia. Retrieved from http://hdl.handle.net/2429/43553

Hodges, N. J., Augaitis, L., & Crocker, P. R. E. (2014). *Sport commitment and deliberate practice among male and female triathletes*. Manuscript under revision.

Hodges, N. J., Kerr, T., Starkes, J. L., & Weir, P. L. (2004). Predicting performance times from deliberate practice hours for triathletes and swimmers: What, when, and where is practice important? *Journal of Experimental Psychology: Applied, 10*, 219–237. doi:10.1037/1076-898X.10.4.219

Hodges, N. J., & Starkes, J. L. (1996). Wrestling with the nature of expertise: A sport-specific test of Ericsson, Krampe, and Tesh-Römer's (1993) theory of deliberate practice. *International Journal of Sport Psychology, 27*, 400–424.

Howard, R. W. (2012). Longitudinal effects of different types of practice on the development of chess expertise. *Applied Cognitive Psychology, 26*, 359–369. doi:10.1002/acp.1834

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. doi:10.1080/13598139.2013.780966

Jabusch, H.-C., Alpers, H., Kopiez, R., Vauth, H., & Altenmuller, E. (2009). The influence of practice on the development of motorskills in pianists: A longitudinal study in a selected motor task. *Human Movement Science, 28*, 74–84. doi:10.1016/j.humov.2008.08.001

Jabusch, H.-C., Yong, R., & Altenmuller, E. (2007). Biographical predictors of music-related motor skills in children pianists. In A. Williamon & D. Coimbra (Eds.), *Proceedings of the International Symposium on Performance Science*. Utrecht, The Netherlands: The European Association of Conservatoires (AEC).

Johnson, M. B., Tenenbaum, G., & Edmonds, W. A. (2006). Adaptation to physically and emotionally demanding conditions: The role of deliberate practice. *High Ability Studies, 17*, 117–136. doi:10.1080/13598130600947184

Kamp, R. J., Dolmans, D. H., van Berkel, H. J., & Schmidt, H. G. (2012). The relationship between students' small group activities, time spent on self-study, and achievement. *Higher Education, 64*, 385–397. doi:10.1007/s10734-011-9500-5

Kopiez, R., Jabusch, H.-C., Galley, N., Homann, J.-C., Lehmann, A. C., & Altenmuller, E. (2011). No disadvantage for left-handed musicians: The relationship between handedness, perceived constraints and performance-related skills in string players and pianists. *Psychology of Music, 40*, 357–384. doi:10.1177/0305735610394708

Kopiez, R., & Lee, J. I. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research, 10*, 41–62. doi: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, 25*, 331–359. doi:10.1037/0096-3445.125.4.331

Law, M. P., Cote, J., & Ericsson, K. A. (2007). Characteristics of expert development in rhythmic gymnastics: A retrospective study. *International Journal of Sport and Exercise Psychology, 5*, 82–103. doi:10.1080/1612197X.2008.9671814 [This effect size was excluded from the meta-analysis; see Open Data at https://osf.io/rhfsk.]

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

Limanond, T., Jomnonkwao, S., Watthanaklang, D., Ratanavaraha, V., & Siridhara, S. (2011). How vehicle ownership affect time utilization on study, leisure, social activities, and academic performance of university students? A case study of engineering freshmen in a rural university in Thailand. *Transport Policy, 18*, 719–726. doi:10.1016/j.tranpol.2011.01.007

Lopez, E. J., Nandagopal, K., Shavelson, R. J., Szu, E., & Penn, J. (2013). Self-regulated learning study strategies and academic performance in undergraduate organic chemistry:

An investigation examining ethnically diverse students. *Journal of Research in Science Teaching, 50*, 660–676. doi:10.1002/tea.21095

Loyens, S. M. M., Gijbels, D., Coertjens, L., & Cote, D. J. (2013). Students' approaches to learning in problem-based learning: Taking into account professional behavior in the tutorial groups, self-study time, and different assessment aspects. *Studies in Educational Evaluation, 39*, 23–32. doi:10.1016/j.stueduc.2012.10.004

Loyens, S. M. M., Rikers, R. M. J. P., & Schmidt, H. G. (2007). The impact of students' conceptions of constructivist assumptions on academic achievement and drop-out. *Studies in Higher Education, 32*, 581–602. doi:10.1080/03075070701573765

Masui, C., Broeckmans, J., Doumen, S., Groenen, A., & Molenberghs, G. (2012). Do diligent students perform better? Complex relations between student and course characteristics, study time, and academic performance in higher education. *Studies in Higher Education, 2012*, 1–23. doi:10.1080/03075079.2012.721350

Maynard, D., Hambrick, D. Z., & Meinz, E. J. (2014). *Practice vs. play as predictors of individual differences in bowling skill*. Unpublished data.

McKinney, E. H., & Davis, K. J. (2003). Effects of deliberate practice on crisis decision performance. *Human Factors, 45*, 436–444. doi:10.1518/hfes.45.3.436.27251

McPherson, G. E. (2005). From child to musician: Skill development during the beginning stages of learning an instrument. *Psychology of Music, 33*, 5–35. doi:10.1177/0305735605048012

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

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

Memmert, D., Baker, J., & Bertsch, C. (2010). Play and practice in the development of sport-specific creativity in team ball sports. *High Ability Studies, 21*, 3–18. doi:10.1080/13598139.2010.488083

Moesch, K., Elbe, A.-M., Hauge, M.-L. T., & Wikman, J. M. (2011). Late specialization: The key to success in centimeters, grams, or seconds (cgs) sports. *Scandinavian Journal of Medicine & Science in Sports, 21*, e282–e290. doi:10.1111/j.1600-0838.2010.01280.x

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30*, 96–116. doi:10.1016/j.cedpsych.2004.06.001

Rosário, P., Núñez, J. C., Valle, A., González-Pienda, J., & Lourenço, A. (2013). Grade level, study time, and grade retention and their effects on motivation, self-regulated learning strategies, and mathematics achievement: A structural equation model. *European Journal of Psychology of Education, 28*, 1311–1331. doi:10.1007/s10212-012-0167-9

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

Ryabov, I. (2012). The effect of time online on grades in online sociology courses. *Journal of Online Learning and Teaching, 8*, 1–12.

Schaap, L., Schmidt, H. G., & Verkoeijen, P. P. J. L. (2012). Assessing knowledge growth in a psychology curriculum: Which students improve most? *Assessment & Evaluation in Higher Education, 37*, 875–887. doi:10.1080/02602938.2011.581747

Schultetus, S., & Charness, N. (1997). *Fencing data*. Unpublished data. Cited in Deakin, J. M., & Cobley, S. (2003). A search for deliberate practice: Expert performance in sports: Advances in research on sport expertise. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 115–136). Champaign, IL: Human Kinetics.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology, 87*, 287–309. doi:10.1111/j.2044-8295.1996.tb02591.x

Smith, A. (2012). *Retrospective practice histories of Division I and Division II female basketball players in the Carolinas* (Unpublished dissertation). University of South Carolina, Columbia. Available from ProQuest Dissertations and Theses database. (UMI No. 3509548)

Snelling, P. C., Lipscomb, M., Lockyer, L., Yates, S., & Young, P. (2010). Time spent studying on a pre-registration nursing programme module: An exploratory study and implications for regulation. *Nurse Education Today, 30*, 713–719. doi:10.1016/j.nedt.2010.01.010

Sonnentag, S., & Kleine, B. M. (2000). Deliberate practice at work: A study with insurance agents. *Journal of Occupational and Organizational Psychology, 73*, 87–102. doi:10.1348/096317900166895

Svanum, S., & Bigatti, S. M. (2006). The influences of course effort and outside activities on grades in a college course. *Journal of College Student Development, 47*, 564–576. doi:10.1353/csd.2006.0063

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

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

Urena, C. A. (2004). *Skill acquisition in ballet dancers: The relationship between deliberate practice and expertise* (Doctoral dissertation). The Florida State University, Tallahassee. Available from Electronic Theses, Treatises and Dissertations, Diginole Commons. (Paper No. 1452)

van Gelder, T., Bissett, M., & Cumming, G. (2004). Cultivating expertise in informal reasoning. *Canadian Journal of Experimental Psychology, 58*, 142–152. doi:10.1037/h0085794

Walker, C., Fleischer, S., & Winn, S. (2008). A path analysis of first-year social science students' engagement with their degree and Level 1 academic outcome. *ELiSS, 1*, 1–19. doi:10.11120/elss.2008.01020003

Ward, P., Hodges, N. J., Starkes, J. L., & Williams, M. A. (2007). The road to excellence: Deliberate practice and the development of expertise. *High Ability Studies, 18*, 119–153. doi:10.1080/13598130701709715

Weissensteiner, J., Abernethy, B., Farrow, D., & Müller, S. (2008). The development of anticipation: A cross-sectional examination of the practice experiences contributing to skill in cricket batting. *Journal of Sport & Exercise Psychology, 30*, 663–684.

Wood, L. J. (1999). *Practice, ability, and expertise in computer programming* (Unpublished doctoral dissertation). University of Minnesota, Minneapolis.

Woody, R. H. (2003). Explaining expressive performance: Component cognitive skills in an aural modeling task. *Journal of Research in Music Education, 51*, 51–63. doi:10.2307/3345648

Young, B. W. (1998). *Deliberate practice and the acquisition of expert performance in Canadian middle-distance running* (Master's thesis). University of Ottawa. Available from Masters Abstracts International. (URI No. 10393/8607)

Young, B. W., Medic, N., Weir, P. L., & Starkes, J. L. (2008). Explaining performance in elite middle-aged runners: Contributions from age and from ongoing and past training factors. *Journal of Sport & Exercise Psychology, 30*, 737–754.

Yu, D. D. (2011). How much do study habits, skills, and attitudes affect student performance in introductory college accounting courses? *New Horizons in Education, 59*, 1–15.

Zimmerman, B. J., & Kitsantas, A. (2005). Homework practices and academic achievement: The mediating role of self-efficacy and perceived responsibility beliefs. *Contemporary Educational Psychology, 30*, 397–417. doi:10.1016/j.cedpsych.2005.05.003

Methodological Details

Effect size calculations

DOI: 10.1177/0956797614535810

- We performed any required conversions from reported statistics to standardized mean differences (Cohen's *d*s; which were then converted into biserial correlations) using the calculator companion to Lipsey and Wilson (2001), found at: http://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD1.php.

Sampling error variance calculations

- We calculated the sampling error variance of each correlation coefficient using the formula of $(1 - r^2)^2/(N - 1)$, which is a standard practice in meta-analysis (Hunter & Schmidt, 2004; Schmidt, Hunter, & Raju, 1998).

Multiple effect size calculations

- If a correlation coefficient was reported, we used this as the effect size. If only categorical comparisons (e.g., professional vs. amateur) were reported, we calculated Cohen's *d* from the means and standard deviations if the study compared two groups and reported means and standard deviations; otherwise, we calculated Cohen's *d* from an *F* statistic or used the reported Cohen's *d*, and then converted the effect size to a biserial correlation coefficient.

Multiple analyses

- When multiple effects were based on a series of data-preparation steps and transformations (e.g., removing outliers, log-transforming deliberate practice), we used the final effect size.

Multiple time points

- When effects were reported for one or more activities interpretable as deliberate practice (henceforth referred to as *deliberate practice*) across a series of time points, we used the effect size from the last time point, reflecting the longest accumulated time of deliberate practice for the sample. In one case (Young, 1998), due to attrition, there were no subjects remaining in one of the skill groups and so we used the longest time in which all groups were represented (the second-to-last time point).

Multiple measures of deliberate practice

- In cases where multiple measures of deliberate practice from a single sample were reported without an overall estimate (e.g., for triathlon performance, deliberate practice for swimming, biking, and running, but no report of total triathlon deliberate practice), we averaged the reported effect sizes and corrected variance by calculating the pooled variance and dividing by the square root of the number of effect sizes.
- When measures of deliberate practice with varying specificity were reported, we used the measure at the level of area (column G in the Open Data file). For example, if a study investigated the effect of deliberate practice on performance in field hockey, we considered field hockey to be the area and therefore used the effect size for the relationship between deliberate practice in field hockey and field hockey performance and not the effect size for the relationship between deliberate practice across all sports and field hockey performance. As another example, if a study investigated the role of deliberate practice on piano sight-reading performance, we considered piano to be the area and therefore used the effect size for the relationship between deliberate practice in piano and sight-reading performance and not the effect size for the relationship between deliberate practice in sight-reading only and sight-reading performance.

Multiple measures of performance

- We adjusted for dependent samples using the method outlined by Cheung and Chan, (2004; 2008), which statistically lowers the associated sample size due to dependent

DOI: 10.1177/0956797614535810

effects being partially redundant, which reduces the weight of these effect sizes in the meta-analysis so as not to overly contribute to the overall mean effect size. The adjustment is based on the initial sample size, the number of dependent correlations, the degree of inter-dependence among the dependent correlations, and the overall average correlation.

Multiple reports of sample size

- When degrees of freedom clearly indicated that the number of subjects entered into the analysis differed from the total number of subjects reported, we used the number of subjects entered into the analysis as the sample size associated with this effect size. If the analysis categorized participants into groups and we could not know whether the missing subjects were proportionally missing from each group, we calculated the effect size with the reported number of participants, but still used the number indicated by the degrees of freedom as the sample size associated with the effect size.

Negative correlations

- We did not detect anything systematic about the negative correlations in terms of their distribution across domains (games = 0, sports = 5, music = 1, education = 1, and professions = 3), or levels of the other moderator variables: Predictability of task environment (high = 2, medium-high = 1, medium = 1, medium-low = 5, low = 1); Method used to assess deliberate practice (interview = 3, questionnaire = 6, log = 1); and Method used to assess performance (group membership = 3, laboratory measure = 3, expert-rated = 3, objective score = 1).

Categorization of Measures of Solitary Deliberate Practice

DOI: 10.1177/0956797614535810

- We categorized measures of deliberate practice as solitary deliberate practice when the authors of the study explicitly stated that the measure of deliberate practice was designed to estimate amount of practice or study alone (i.e., not with others).

Predictability of the Task Environment

- Predictability of the task environment takes into account 1) the rate of change of the task environment to which a performer must respond, and 2) the range of probable stimuli to which a performer may respond. Each of these components was scored on a scale of 1-3 and the average was taken to produce the predictability score. Invariance: 3 = highly static, 2 = somewhat dynamic, 1 = highly dynamic. Constraints: 3 = highly constrained, 2 = moderately constrained, 1 = low constraints. See Open Data file for further descriptions and scores. See Figure S1 for a scatterplot depicting the linear relationship between predictability scores and effect sizes.

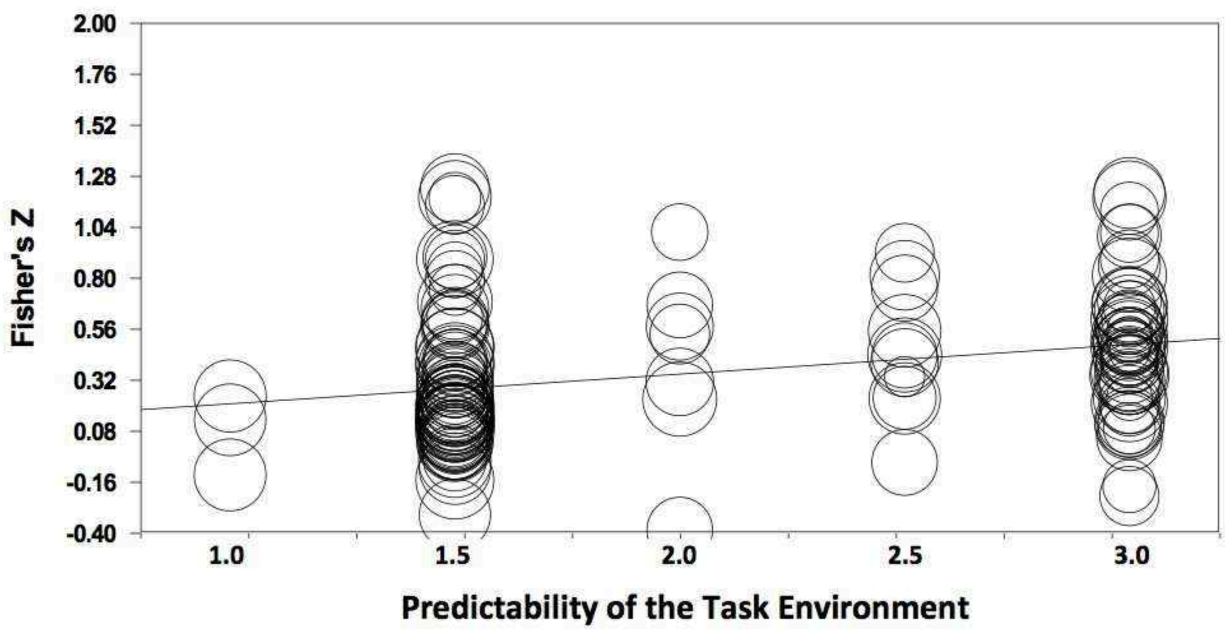


Figure S1. Predictability of the Task Environment as a Moderator of the Relationship Between Deliberate Practice and Performance. Scatterplot representing the relationship between predictability of the task environment and the Fisher's Z transformation of the deliberate practice-performance effect size. Circles represent effect sizes. The size of the circles represents the weight (inverse sampling variance) of the effect size (larger circles = greater weight). The positive slope indicates that the relationship between deliberate practice and performance is stronger (more positive) for tasks higher in predictability than for tasks lower in predictability.

Variable Rating

- In addition to the authors' ratings, two raters independently categorized and scored the moderator variables. Inter-rater reliability for the three ratings of the nominal variables (domain, deliberate practice measure, and performance measure) was calculated using Cohen's Kappa. Inter-rater reliability for the three ratings of the two rank order variables, the sub-components of predictability (invariance and constraints), was calculated using Spearman's correlation coefficient. Inter-rater reliability for domain was .99-1.0, for invariance was .89-.94, for constraints was .91-.96, for measure of deliberate practice was .91-.98, and for measure of performance was .78-.83.

Reliability Assumptions

- In order to test the relationship between deliberate practice and performance across a range of reliability assumptions, we corrected for a variety of hypothetical measurement unreliability scenarios (see Table S1) using the following formula: $\hat{r} = r_{xy}/(r_{xx}r_{yy})^{1/2}$, where r_{xx} and r_{yy} are reliability coefficients for x and y, respectively (Schmidt & Hunter, 1996). Deliberate practice measures are typically found to have acceptable or better reliability ($\geq$ .70; see Ericsson, 2013; Tuffiash, Roring, & Ericsson, 2007, for specific claims concerning reliability). Reliability estimates of performance measures are typically .80 or higher (Hambrick et al. 2014).

DOI: 10.1177/0956797614535810

DOI: 10.1177/0956797614535810

Table S1

Variance in Performance Explained (and Unexplained) by Deliberate Practice Corrected for Measurement Unreliability Across a Range of Reliability Assumptions

Overall
[uncorrected variance explained = 12% (88%)]

Performance Measure Reliability	Deliberate Practice Measure Reliability			
	.60	.70	.80	.90
.90	22% (78%)	19% (81%)	17% (83%)	15% (85%)
.80	25% (75%)	21% (79%)	19% (81%)	17% (83%)
.70	28% (72%)	24% (76%)	21% (79%)	19% (81%)
.60	33% (67%)	28% (72%)	25% (75%)	22% (78%)

Games
[uncorrected variance explained = 26% (74%)]

Performance Measure Reliability	Deliberate Practice Measure Reliability			
	.60	.70	.80	.90
.90	49% (51%)	42% (58%)	37% (63%)	33% (67%)
.80	55% (45%)	47% (53%)	41% (59%)	37% (63%)
.70	63% (37%)	54% (46%)	47% (53%)	42% (58%)
.60	74% (26%)	63% (37%)	55% (45%)	49% (51%)

Music
[uncorrected variance explained = 21% (79%)]

Performance Measure Reliability	Deliberate Practice Measure Reliability			
	.60	.70	.80	.90
.90	39% (61%)	33% (67%)	29% (71%)	26% (74%)
.80	43% (57%)	37% (63%)	33% (67%)	29% (71%)
.70	50% (50%)	42% (58%)	37% (63%)	33% (67%)
.60	58% (42%)	50% (50%)	43% (57%)	39% (61%)

Sports
[uncorrected variance explained = 18% (82%)]

Performance Measure Reliability	Deliberate Practice Measure Reliability			
	.60	.70	.80	.90
.90	33% (67%)	29% (71%)	25% (75%)	22% (78%)
.80	37% (63%)	32% (68%)	28% (72%)	25% (75%)
.70	43% (57%)	37% (63%)	32% (68%)	29% (71%)
.60	50% (50%)	43% (57%)	37% (63%)	33% (67%)

DOI: 10.1177/0956797614535810

Education
[uncorrected variance explained = 4% (96%)]

Performance Measure Reliability	Deliberate Practice Measure Reliability			
	.60	.70	.80	.90
.90	8% (92%)	7% (93%)	6% (94%)	5% (95%)
.80	9% (91%)	8% (92%)	7% (93%)	6% (94%)
.70	10% (90%)	9% (91%)	8% (92%)	7% (93%)
.60	12% (88%)	10% (90%)	9% (91%)	8% (92%)

Professions
[uncorrected variance explained < 1% (> 99%)]

Performance Measure Reliability	Deliberate Practice Measure Reliability			
	.60	.70	.80	.90
.90	<1% (>99%)	<1% (>99%)	<1% (>99%)	<1% (>99%)
.80	1% (99%)	<1% (>99%)	<1% (>99%)	<1% (>99%)
.70	1% (99%)	<1% (>99%)	<1% (>99%)	<1% (>99%)
.60	1% (99%)	1% (99%)	1% (99%)	<1% (>99%)

Additional Publication Bias Analyses

Publication Bias Analyses

- Publication bias occurs when weak effect sizes are systematically suppressed from publication. Funnel plots depict the relationship between sample size and effect size and are inspected for asymmetry, that is, if relatively small studies with relatively weak effect sizes—studies prone to remain unpublished—are missing from the meta-analysis, the plot will be asymmetrical with fewer effect sizes on the left side of the plot. Figure S2 shows the funnel plot for the present meta-analysis. This funnel plot is approximately symmetrical suggesting that publication bias is not affecting the results of our meta-analysis.

DOI: 10.1177/0956797614535810

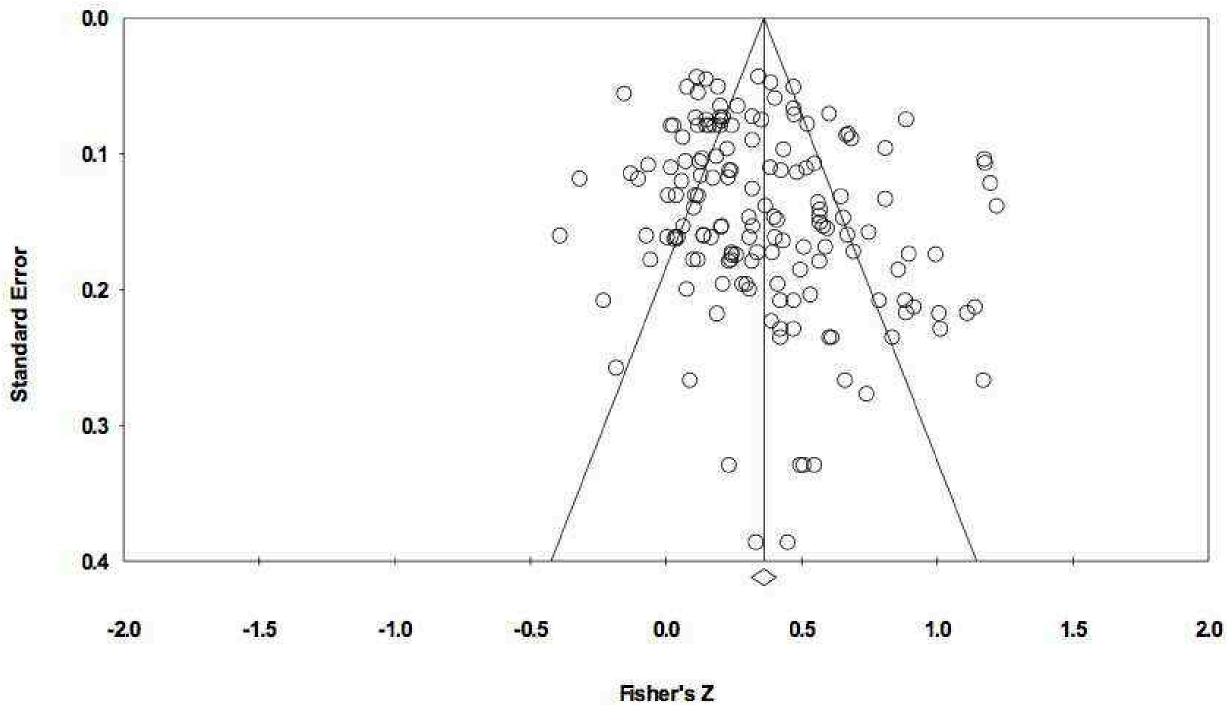


Figure S2. Funnel plot. Circles represent effect sizes included in the meta-analysis. Effect sizes are plotted by size (Fisher's *Z* correlation coefficient; *x* axis) and standard error (calculated from sample size; *y* axis).

- We also performed publication bias analyses by domain. Inspection of funnel plots and trim-and-fill analyses (Duval & Tweedie, 2000a, 2000b) suggested that no weaker-than-average effect sizes were missing from the meta-analysis overall or by domain.
- We also tested for publication bias empirically, that is, by entering publication status (published, unpublished) as a moderator variable and testing differences in effect sizes between the groups. Overall, no significant difference emerged between effect sizes from published studies ($\bar{r}$ = .35) and unpublished studies ($\bar{r}$ = .32), Q(1) = .23, *p* = .63. All effect sizes from the domain of Games were published. Only one effect size each from the domains of Music, Education, and Professions was unpublished. In the domain of Sports, there was a significant difference between effect sizes from published studies, (34

DOI: 10.1177/0956797614535810

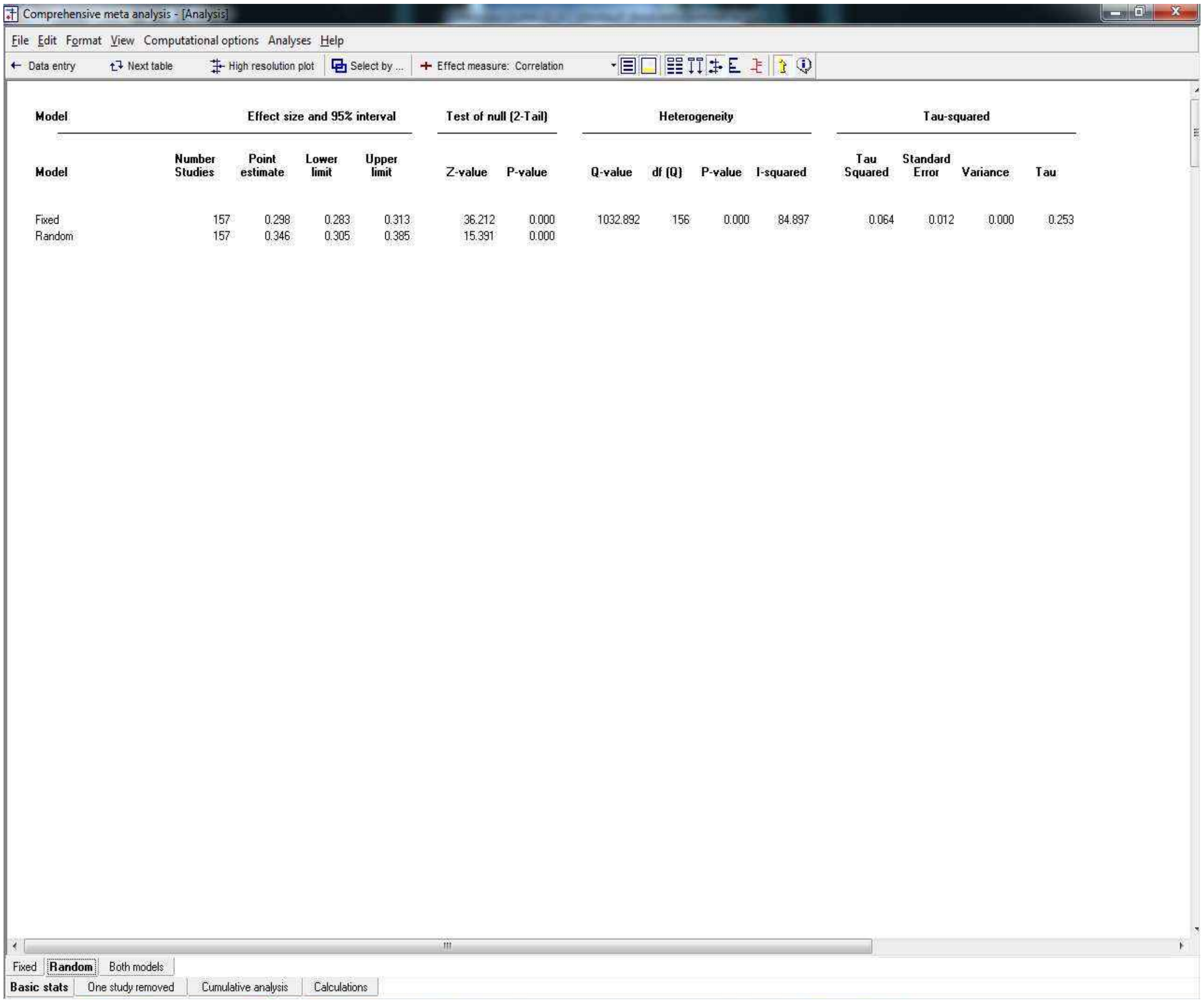
cases, $\bar{r}$ = .50, 25% of the variance explained) and unpublished studies (26 cases, $\bar{r}$ = .32, 10% of the variance explained), Q(1) = 5.45, p = .02.

III. Screen Shots of Results

Screen Shots

- Next, we provide the screen shots of the results from the meta-analysis statistical software used to conduct this meta-analysis (Comprehensive Meta-Analysis; Borenstein, Hedges, Higgins, & Rothstein, 2005). We used random-effects models for analyzing overall effects and mixed-effects models for analyzing the effects of moderator variables. Note: while values on the screen shots generally round to three decimal places, the software user can view the results to 15 decimal places when copying the information from the statistical software and pasting to spreadsheet software such as Excel. We rounded to two decimal places from the full 15-decimal place results.

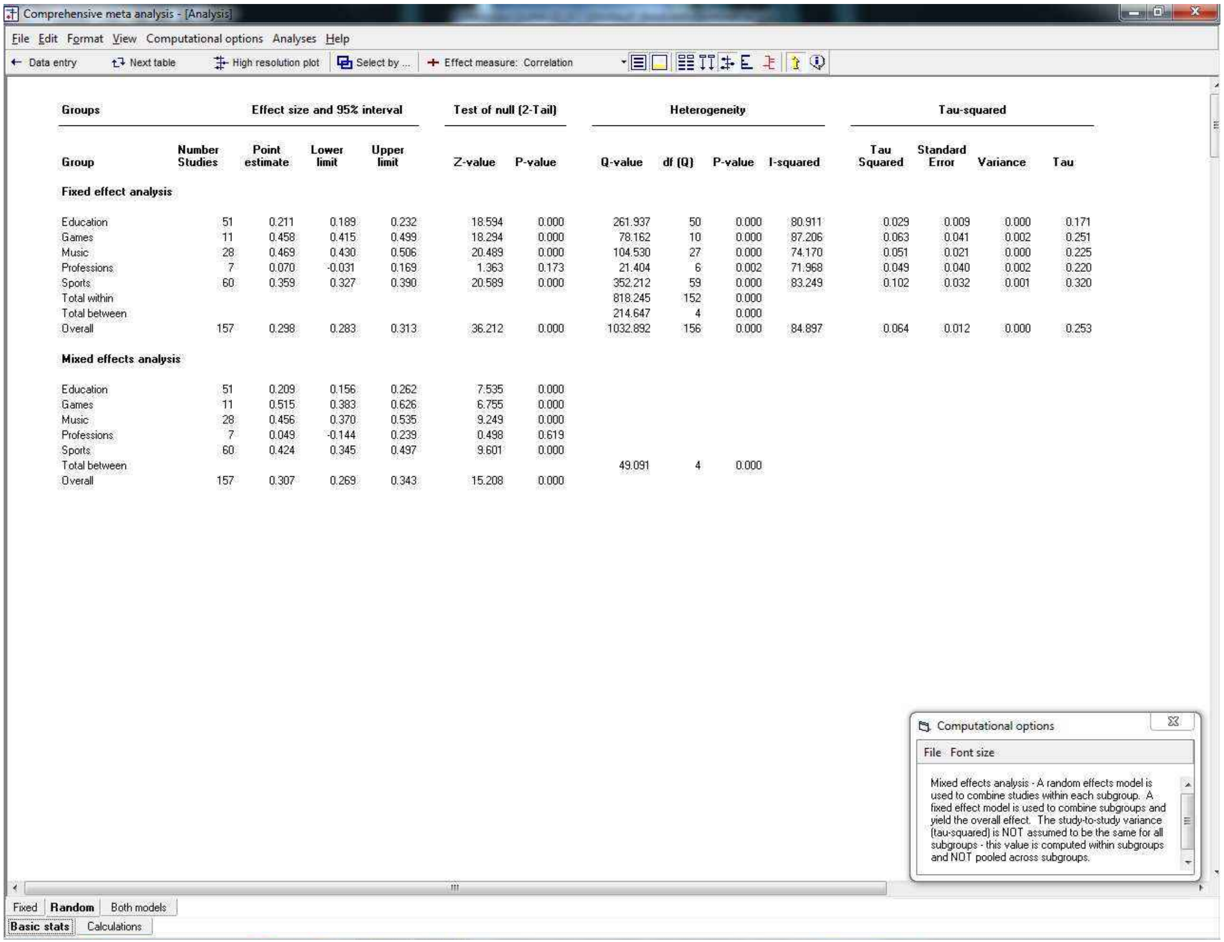
DOI: 10.1177/0956797614535810



Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	157	0.298	0.283	0.313	36.212	0.000	1032.892	156	0.000	84.897	0.064	0.012	0.000	0.253
Random	157	0.346	0.305	0.385	15.391	0.000								

Figure S3. Screen Shot of the Overall Results (Main Model: All Effect Sizes). Note: we report the lower limit of the 95% confidence interval of the point estimate as .30 (not .31) because we rounded to two decimals based on the number in the third decimal place. The lower limit of the 95% confidence interval of the point estimate was .3046. Thus, we rounded this number to .30 and the software rounded this number to .305 for the screen display.

DOI: 10.1177/0956797614535810



Comprehensive meta analysis - [Analysis]

File Edit Format View Computational options Analyses Help

Data entry | Next table | High resolution plot | Select by ... | Effect measure: Correlation

Groups		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
Education	51	0.211	0.189	0.232	18.594	0.000	261.937	50	0.000	80.911	0.029	0.009	0.000	0.171
Games	11	0.458	0.415	0.499	18.294	0.000	78.162	10	0.000	87.206	0.063	0.041	0.002	0.251
Music	28	0.469	0.430	0.506	20.489	0.000	104.530	27	0.000	74.170	0.051	0.021	0.000	0.225
Professions	7	0.070	-0.031	0.169	1.363	0.173	21.404	6	0.002	71.968	0.049	0.040	0.002	0.220
Sports	60	0.359	0.327	0.390	20.589	0.000	352.212	59	0.000	83.249	0.102	0.032	0.001	0.320
Total within							818.245	152	0.000					
Total between							214.647	4	0.000					
Overall	157	0.298	0.283	0.313	36.212	0.000	1032.892	156	0.000	84.897	0.064	0.012	0.000	0.253
Mixed effects analysis														
Education	51	0.209	0.156	0.262	7.535	0.000								
Games	11	0.515	0.383	0.626	6.755	0.000								
Music	28	0.456	0.370	0.535	9.249	0.000								
Professions	7	0.049	-0.144	0.239	0.498	0.619								
Sports	60	0.424	0.345	0.497	9.601	0.000								
Total between							49.091	4	0.000					
Overall	157	0.307	0.269	0.343	15.208	0.000								

Computational options

File Font size

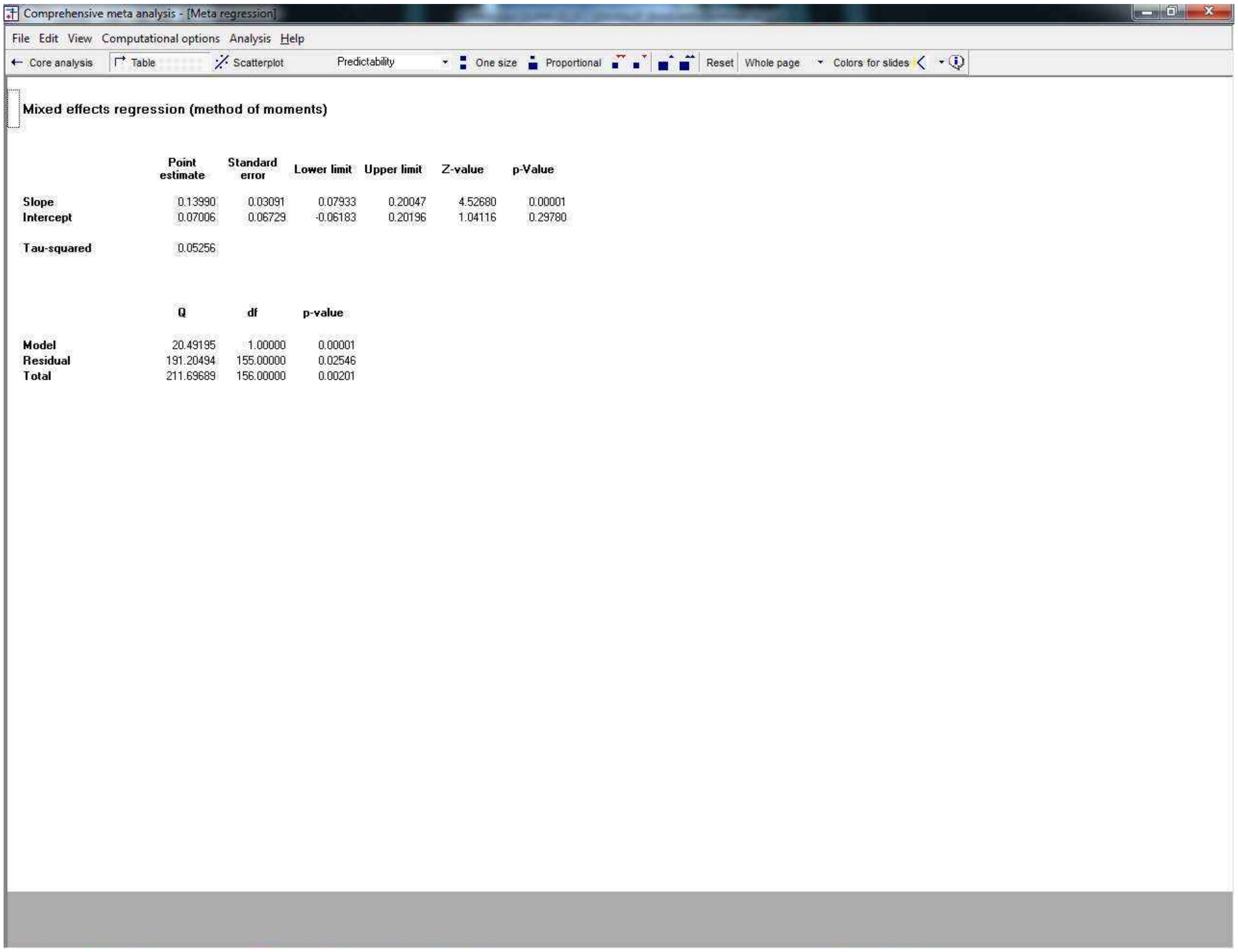
Mixed effects analysis - A random effects model is used to combine studies within each subgroup. A fixed effect model is used to combine subgroups and yield the overall effect. The study-to-study variance (tau-squared) is NOT assumed to be the same for all subgroups - this value is computed within subgroups and NOT pooled across subgroups.

Fixed | Random | Both models

Basic stats | Calculations

Figure S4. Screen Shot of the Results with Domain as a Moderator Variable (Main Model: All Effect Sizes). Note: we report the mean correlation coefficient (point estimate) for Games as .51 (not .52) because we rounded all point estimates to two decimals based on the number in the third decimal place. For Games, the point estimate was .5147. Thus, we rounded this number to .51 and the software rounded this number to .515 for the screen display.

DOI: 10.1177/0956797614535810


Comprehensive meta analysis - [Meta regression]

File Edit View Computational options Analysis Help

Core analysis | Table | Scatterplot | Predictability | One size | Proportional | Reset | Whole page | Colors for slides

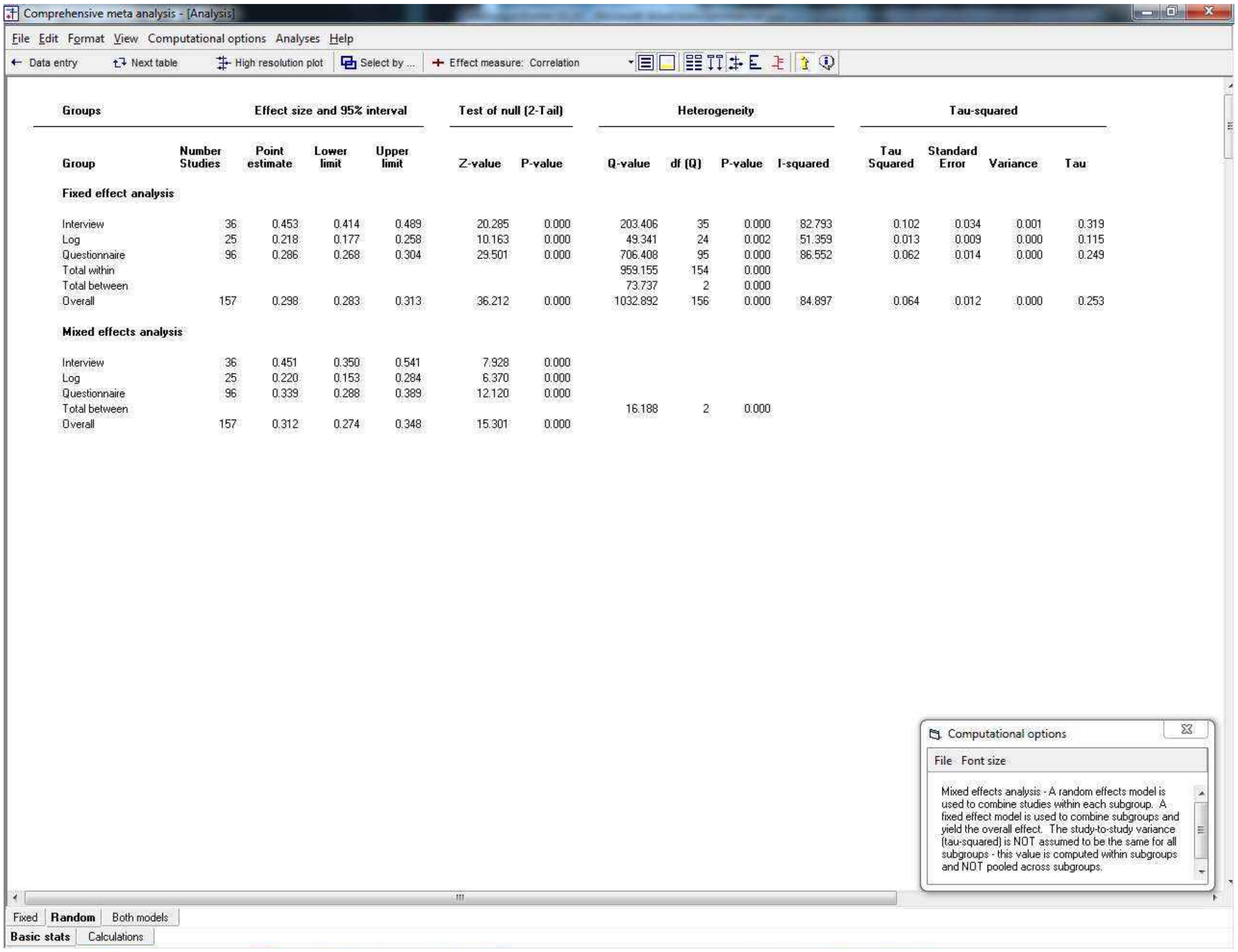
Mixed effects regression (method of moments)

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.13990	0.03091	0.07933	0.20047	4.52680	0.00001
Intercept	0.07006	0.06729	-0.06183	0.20196	1.04116	0.29780
Tau-squared	0.05256					

	Q	df	p-value
Model	20.49195	1.00000	0.00001
Residual	191.20494	155.00000	0.02546
Total	211.69689	156.00000	0.00201

Figure S5. Screen Shot of the Results with Predictability of the Task Environment as a Moderator Variable (Main Model: All Effect Sizes).

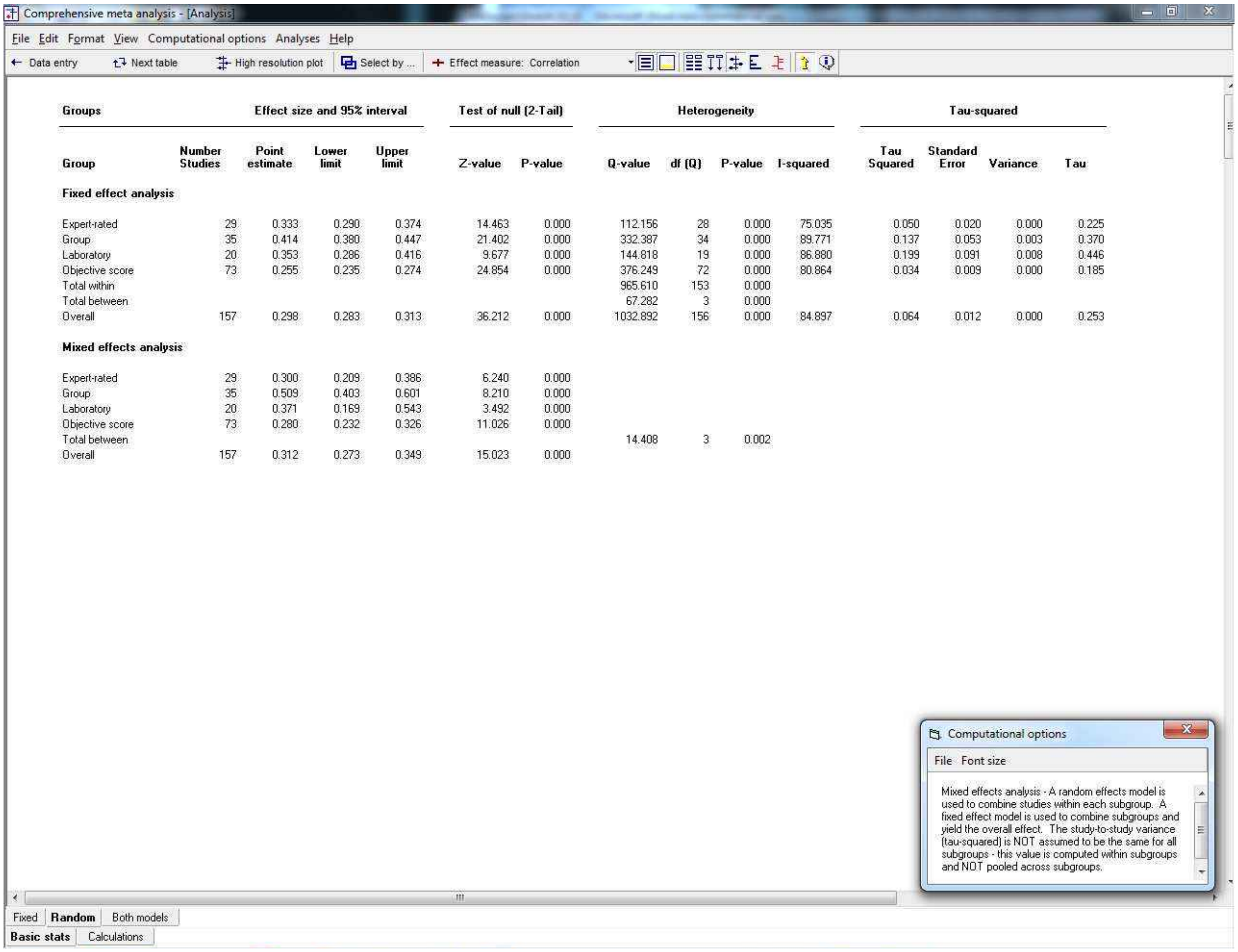
DOI: 10.1177/0956797614535810



Groups	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity				Tau-squared			
Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
Interview	36	0.453	0.414	0.489	20.285	0.000	203.406	35	0.000	82.793	0.102	0.034	0.001	0.319
Log	25	0.218	0.177	0.258	10.163	0.000	49.341	24	0.002	51.359	0.013	0.009	0.000	0.115
Questionnaire	96	0.286	0.268	0.304	29.501	0.000	706.408	95	0.000	86.552	0.062	0.014	0.000	0.249
Total within							959.155	154	0.000					
Total between							73.737	2	0.000					
Overall	157	0.298	0.283	0.313	36.212	0.000	1032.892	156	0.000	84.897	0.064	0.012	0.000	0.253
Mixed effects analysis														
Interview	36	0.451	0.350	0.541	7.928	0.000								
Log	25	0.220	0.153	0.284	6.370	0.000								
Questionnaire	96	0.339	0.288	0.389	12.120	0.000								
Total between							16.188	2	0.000					
Overall	157	0.312	0.274	0.348	15.301	0.000								

Figure S6. Screen Shot of the Results with Method Used to Assess Deliberate Practice as a Moderator Variable (Main Model: All Effect Sizes).

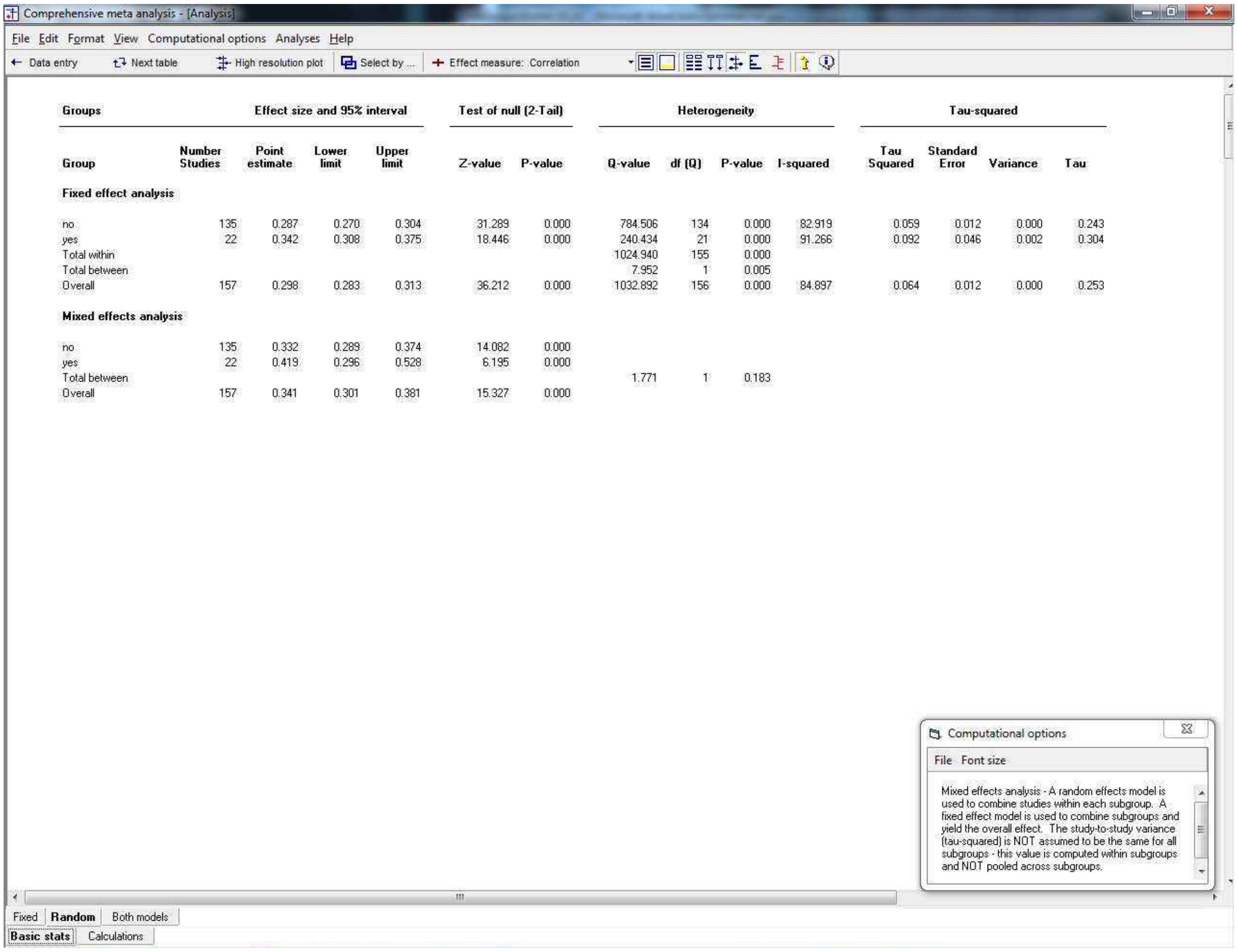
DOI: 10.1177/0956797614535810



Groups		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
Expert-rated	29	0.333	0.290	0.374	14.463	0.000	112.156	28	0.000	75.035	0.050	0.020	0.000	0.225
Group	35	0.414	0.380	0.447	21.402	0.000	332.387	34	0.000	89.771	0.137	0.053	0.003	0.370
Laboratory	20	0.353	0.286	0.416	9.677	0.000	144.818	19	0.000	86.880	0.199	0.091	0.008	0.446
Objective score	73	0.255	0.235	0.274	24.854	0.000	376.249	72	0.000	80.864	0.034	0.009	0.000	0.185
Total within							965.610	153	0.000					
Total between							67.282	3	0.000					
Overall	157	0.298	0.283	0.313	36.212	0.000	1032.892	156	0.000	84.897	0.064	0.012	0.000	0.253
Mixed effects analysis														
Expert-rated	29	0.300	0.209	0.386	6.240	0.000								
Group	35	0.509	0.403	0.601	8.210	0.000								
Laboratory	20	0.371	0.169	0.543	3.492	0.000								
Objective score	73	0.280	0.232	0.326	11.026	0.000								
Total between							14.408	3	0.002					
Overall	157	0.312	0.273	0.349	15.023	0.000								

Figure S7. Screen Shot of the Results with Method Used to Assess Performance as a Moderator Variable (Main Model: All Effect Sizes).

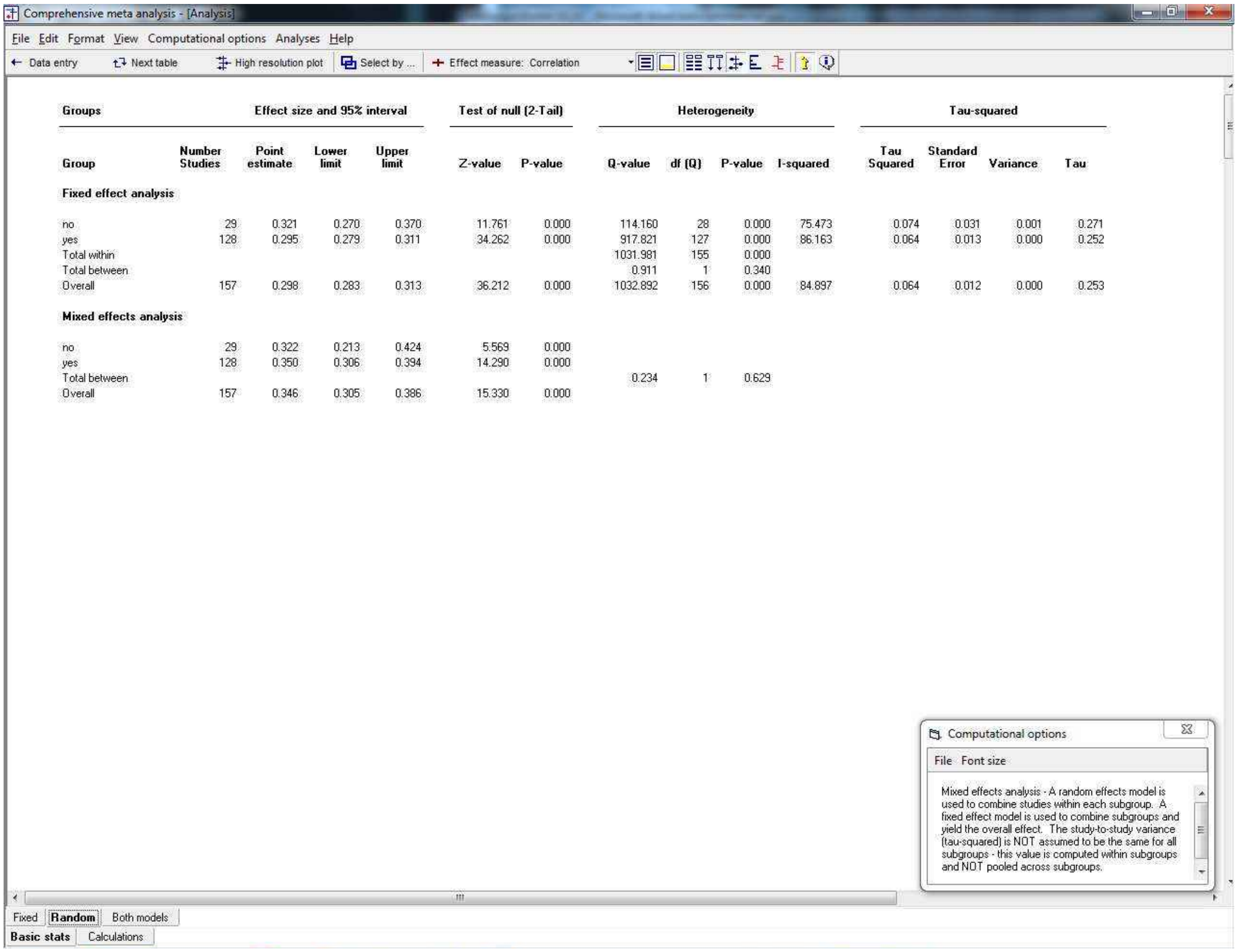
DOI: 10.1177/0956797614535810



Groups		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
no	135	0.287	0.270	0.304	31.289	0.000	784.506	134	0.000	82.919	0.059	0.012	0.000	0.243
yes	22	0.342	0.308	0.375	18.446	0.000	240.434	21	0.000	91.266	0.092	0.046	0.002	0.304
Total within							1024.940	155	0.000					
Total between							7.952	1	0.005					
Overall	157	0.298	0.283	0.313	36.212	0.000	1032.892	156	0.000	84.897	0.064	0.012	0.000	0.253
Mixed effects analysis														
no	135	0.332	0.289	0.374	14.082	0.000								
yes	22	0.419	0.296	0.528	6.195	0.000								
Total between							1.771	1	0.183					
Overall	157	0.341	0.301	0.381	15.327	0.000								

Figure S8. Screen Shot of the Results with Transformed (i.e., Deliberate Practice Log- or Square Root-transformed vs. Untransformed) as a Moderator Variable (Main Model: All Effect Sizes).

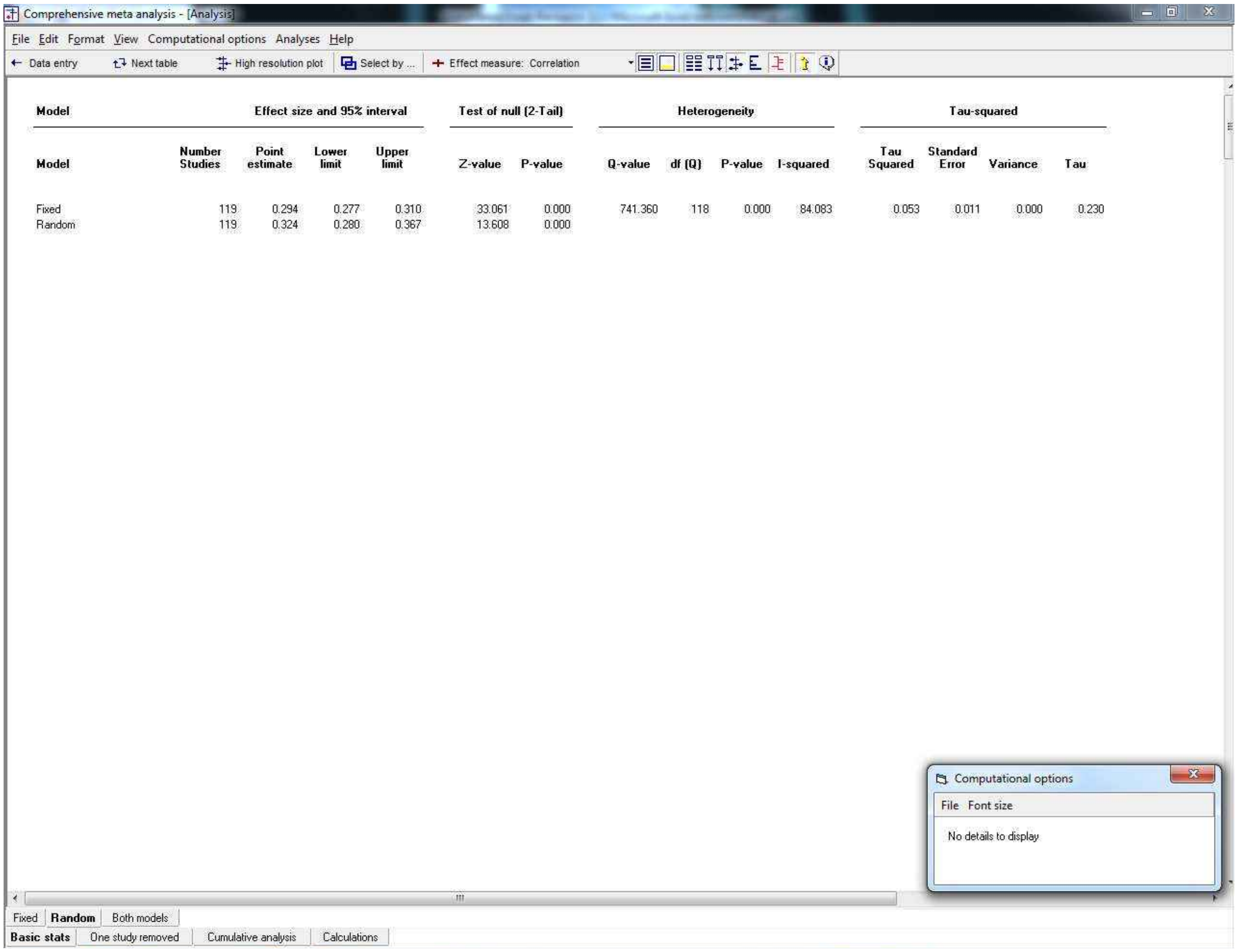
DOI: 10.1177/0956797614535810



Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
no	29	0.321	0.270	0.370	11.761	0.000	114.160	28	0.000	75.473	0.074	0.031	0.001	0.271
yes	128	0.295	0.279	0.311	34.262	0.000	917.821	127	0.000	86.163	0.064	0.013	0.000	0.252
Total within							1031.981	155	0.000					
Total between							0.911	1	0.340					
Overall	157	0.298	0.283	0.313	36.212	0.000	1032.892	156	0.000	84.897	0.064	0.012	0.000	0.253
Mixed effects analysis														
no	29	0.322	0.213	0.424	5.569	0.000								
yes	128	0.350	0.306	0.394	14.290	0.000								
Total between							0.234	1	0.629					
Overall	157	0.346	0.305	0.386	15.330	0.000								

Figure S9. Screen Shot of the Results with Publication Status (i.e., Published vs. Unpublished) as a Moderator Variable (Main Model: All Effect Sizes).

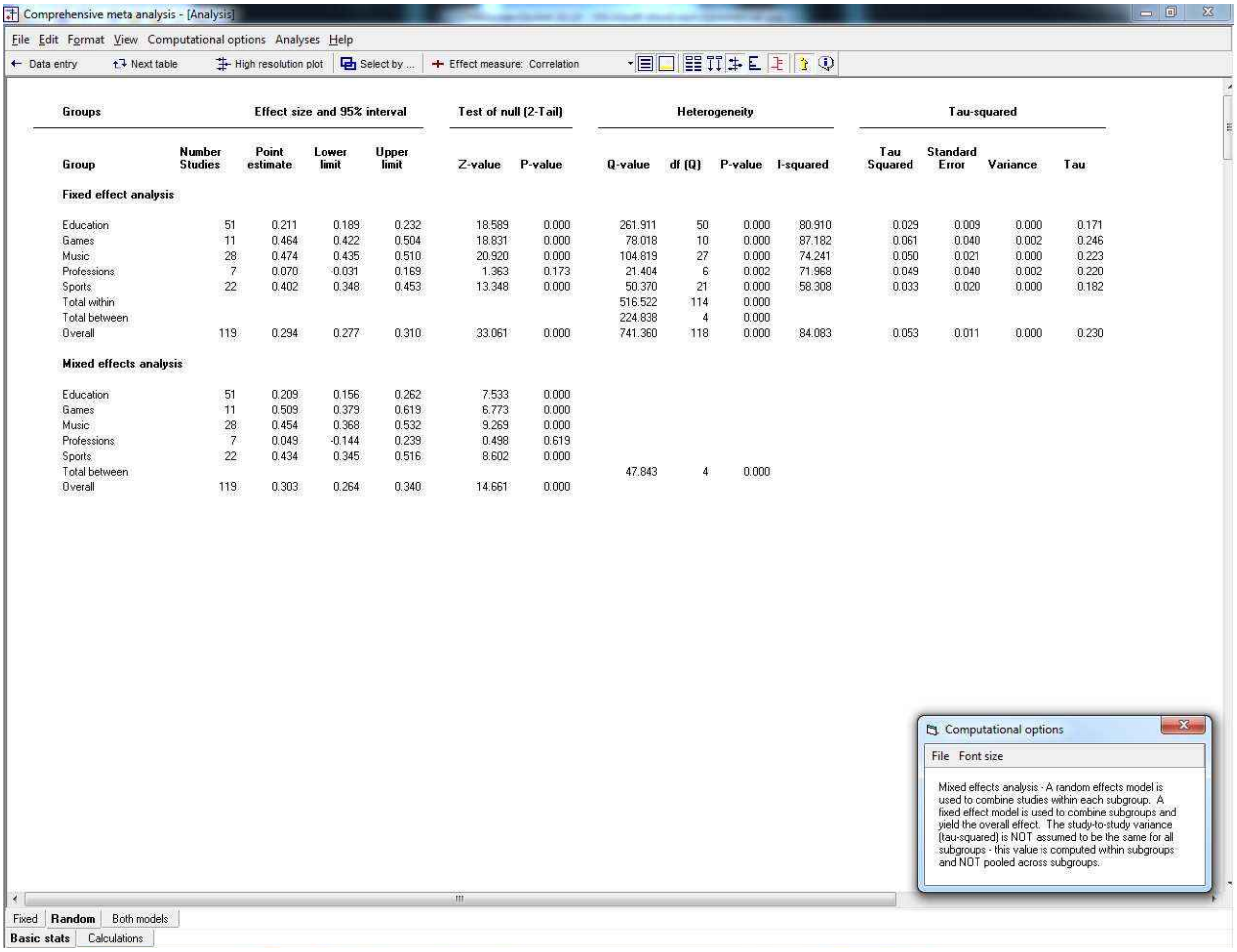
DOI: 10.1177/0956797614535810



Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	119	0.294	0.277	0.310	33.061	0.000	741.360	118	0.000	84.083	0.053	0.011	0.000	0.230
Random	119	0.324	0.280	0.367	13.608	0.000								

Figure S10. Screen Shot of the Overall Results (Additional Model 1: Excluding Team Sports).

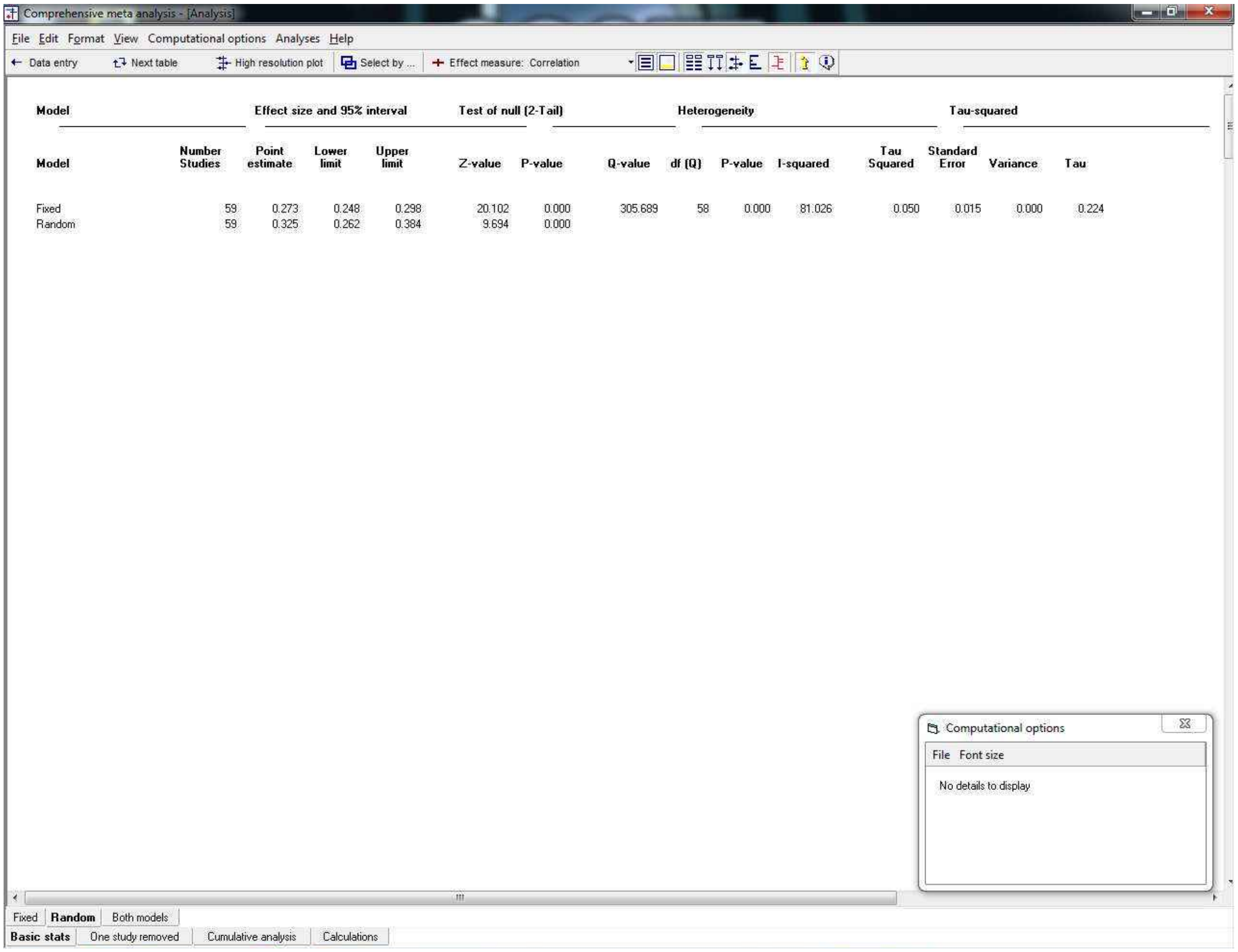
DOI: 10.1177/0956797614535810



Groups	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity				Tau-squared			
Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
Education	51	0.211	0.189	0.232	18.589	0.000	261.911	50	0.000	80.910	0.029	0.009	0.000	0.171
Games	11	0.464	0.422	0.504	18.831	0.000	78.018	10	0.000	87.182	0.061	0.040	0.002	0.246
Music	28	0.474	0.435	0.510	20.920	0.000	104.819	27	0.000	74.241	0.050	0.021	0.000	0.223
Professions	7	0.070	-0.031	0.169	1.363	0.173	21.404	6	0.002	71.968	0.049	0.040	0.002	0.220
Sports	22	0.402	0.348	0.453	13.348	0.000	50.370	21	0.000	58.308	0.033	0.020	0.000	0.182
Total within							516.522	114	0.000					
Total between							224.838	4	0.000					
Overall	119	0.294	0.277	0.310	33.061	0.000	741.360	118	0.000	84.083	0.053	0.011	0.000	0.230
Mixed effects analysis														
Education	51	0.209	0.156	0.262	7.533	0.000								
Games	11	0.509	0.379	0.619	6.773	0.000								
Music	28	0.454	0.368	0.532	9.269	0.000								
Professions	7	0.049	-0.144	0.239	0.498	0.619								
Sports	22	0.434	0.345	0.516	8.602	0.000								
Total between							47.843	4	0.000					
Overall	119	0.303	0.264	0.340	14.661	0.000								

Figure S11. Screen Shot of the Results with Domain as a Moderator Variable (Additional Model 1: Excluding Team Sports).

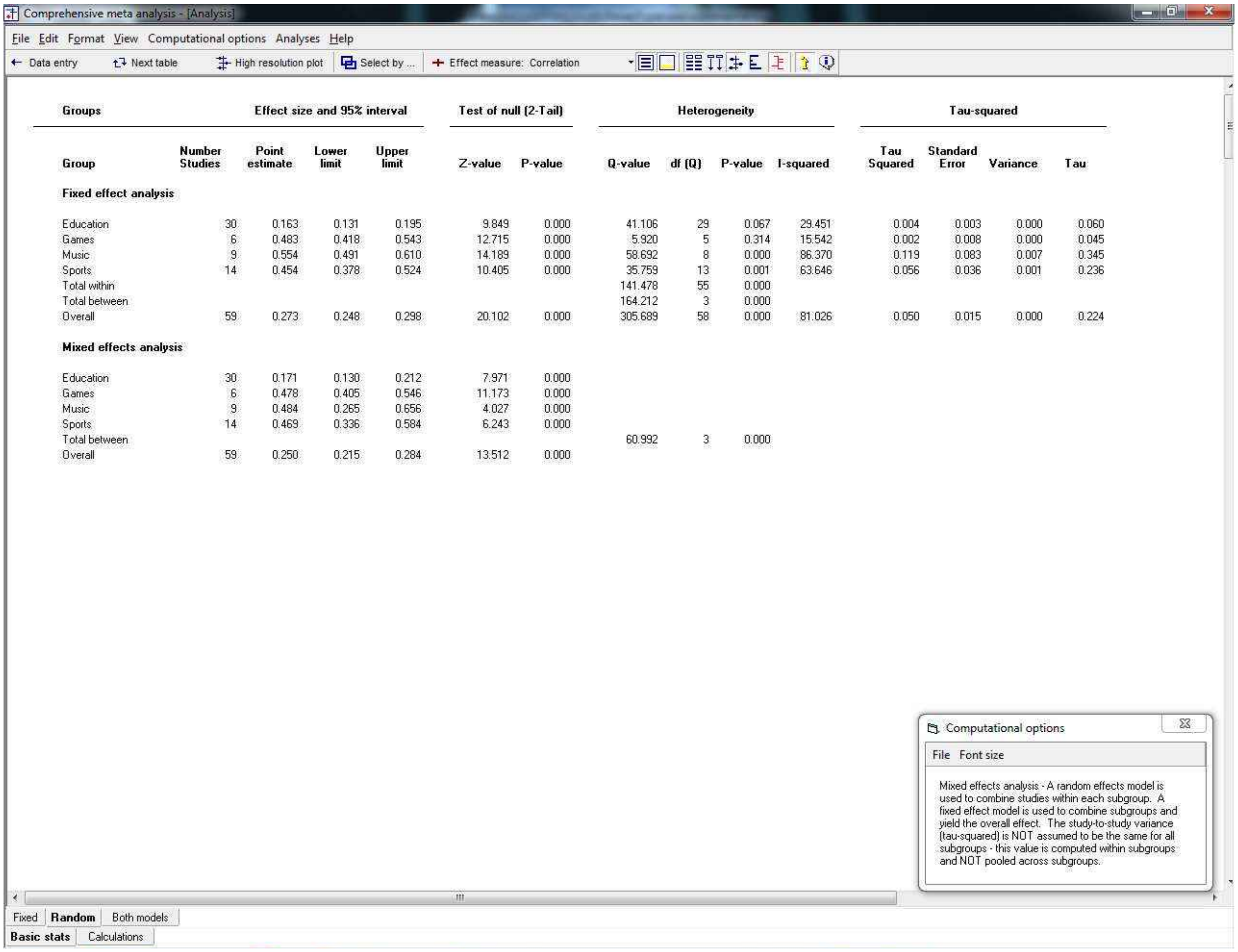
DOI: 10.1177/0956797614535810



Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	59	0.273	0.248	0.298	20.102	0.000	305.689	58	0.000	81.026	0.050	0.015	0.000	0.224
Random	59	0.325	0.262	0.384	9.694	0.000								

Figure S12. Screen Shot of the Overall Results (Additional Model 2: Only Solitary Deliberate Practice).

DOI: 10.1177/0956797614535810



Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
Education	30	0.163	0.131	0.195	9.849	0.000	41.106	29	0.067	29.451	0.004	0.003	0.000	0.060
Games	6	0.483	0.418	0.543	12.715	0.000	5.920	5	0.314	15.542	0.002	0.008	0.000	0.045
Music	9	0.554	0.491	0.610	14.189	0.000	58.692	8	0.000	86.370	0.119	0.083	0.007	0.345
Sports	14	0.454	0.378	0.524	10.405	0.000	35.759	13	0.001	63.646	0.056	0.036	0.001	0.236
Total within							141.478	55	0.000					
Total between							164.212	3	0.000					
Overall	59	0.273	0.248	0.298	20.102	0.000	305.689	58	0.000	81.026	0.050	0.015	0.000	0.224
Mixed effects analysis														
Education	30	0.171	0.130	0.212	7.971	0.000								
Games	6	0.478	0.405	0.546	11.173	0.000								
Music	9	0.484	0.265	0.656	4.027	0.000								
Sports	14	0.469	0.336	0.584	6.243	0.000								
Total between							60.992	3	0.000					
Overall	59	0.250	0.215	0.284	13.512	0.000								

Figure S13. Screen Shot of the Results with Domain as a Moderator Variable (Additional Model 2: Only Solitary Deliberate Practice).

DOI: 10.1177/0956797614535810

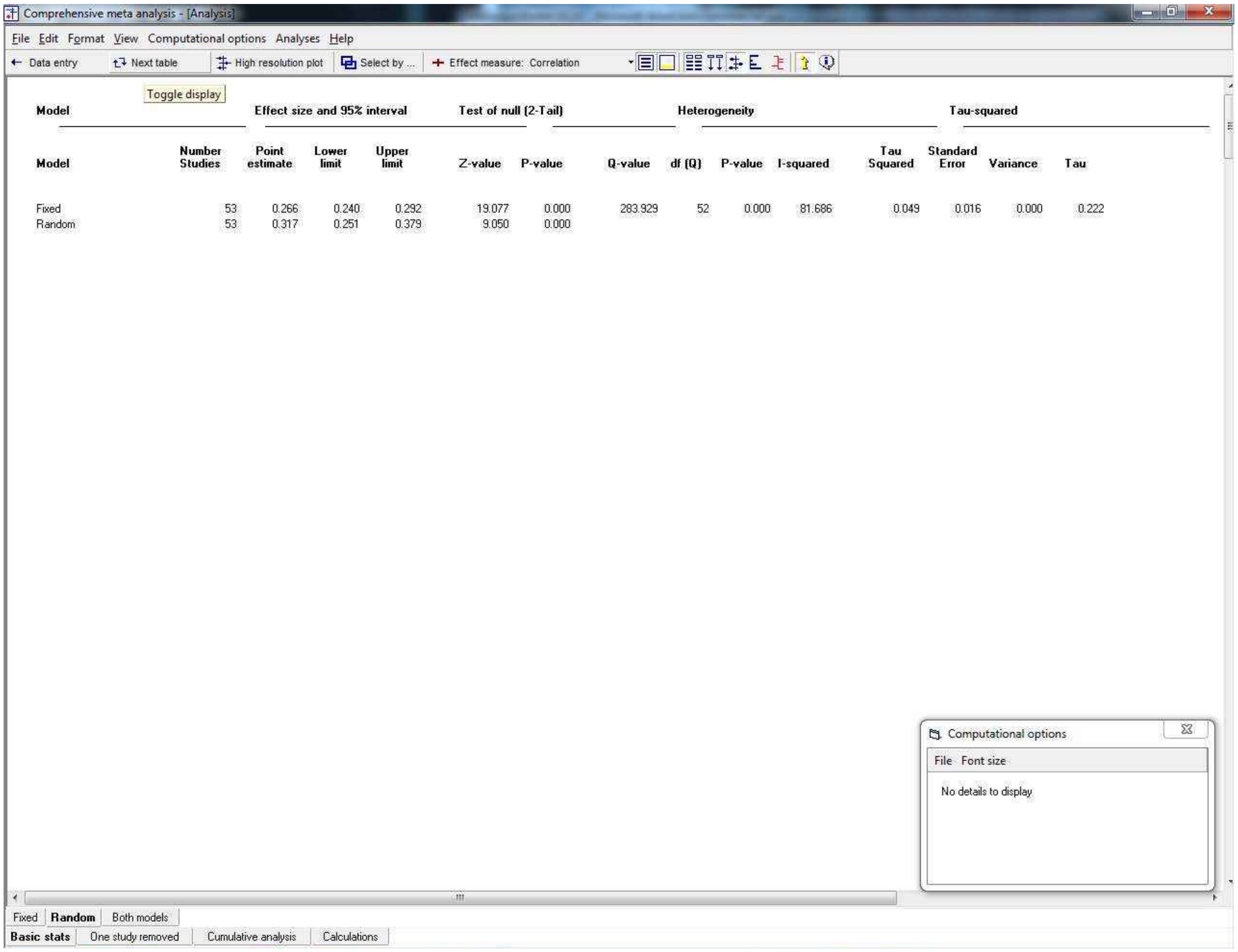
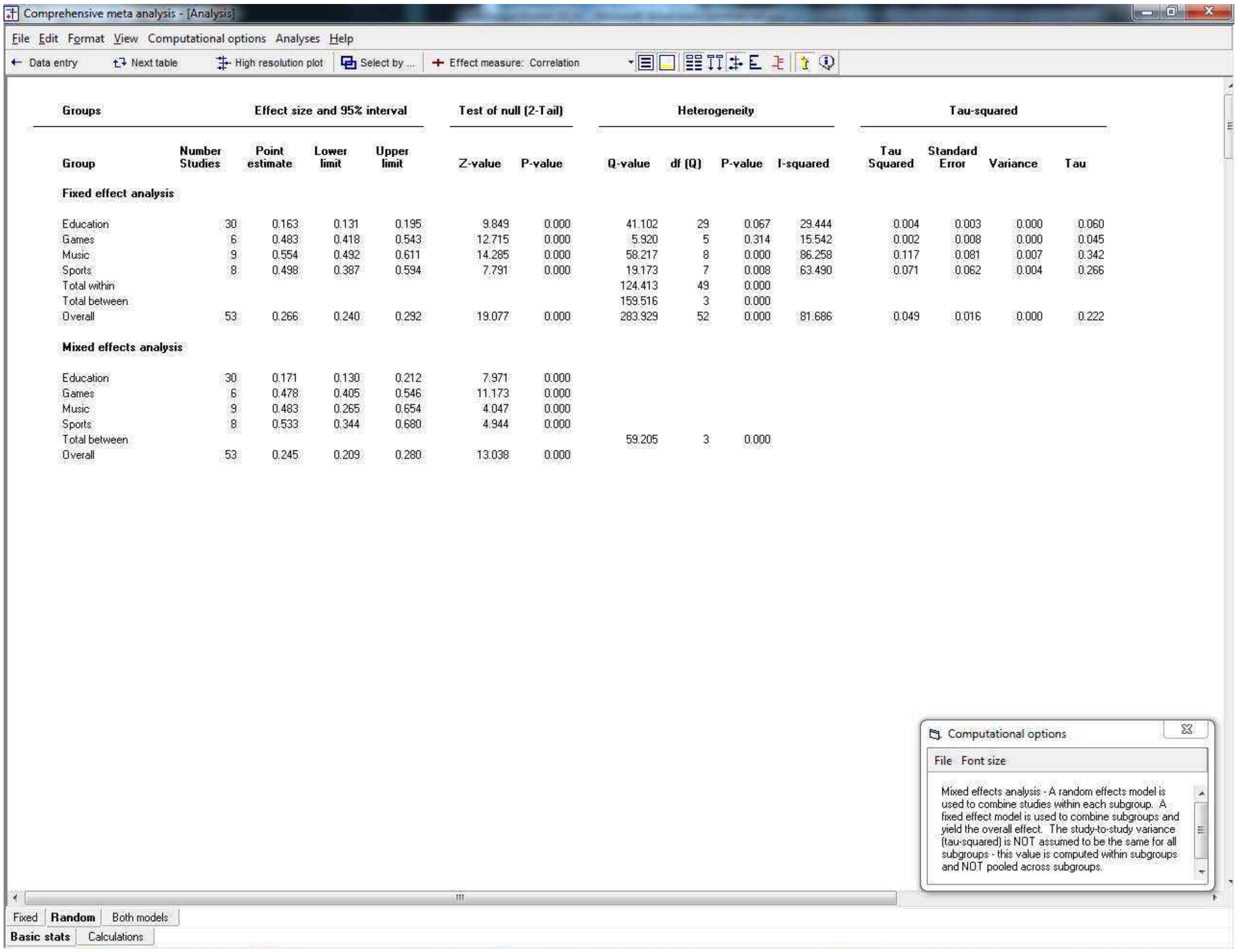


Figure S14. Screen Shot of the Overall Results (Additional Model 3: Excluding Team Sports, Only Solitary Deliberate Practice).

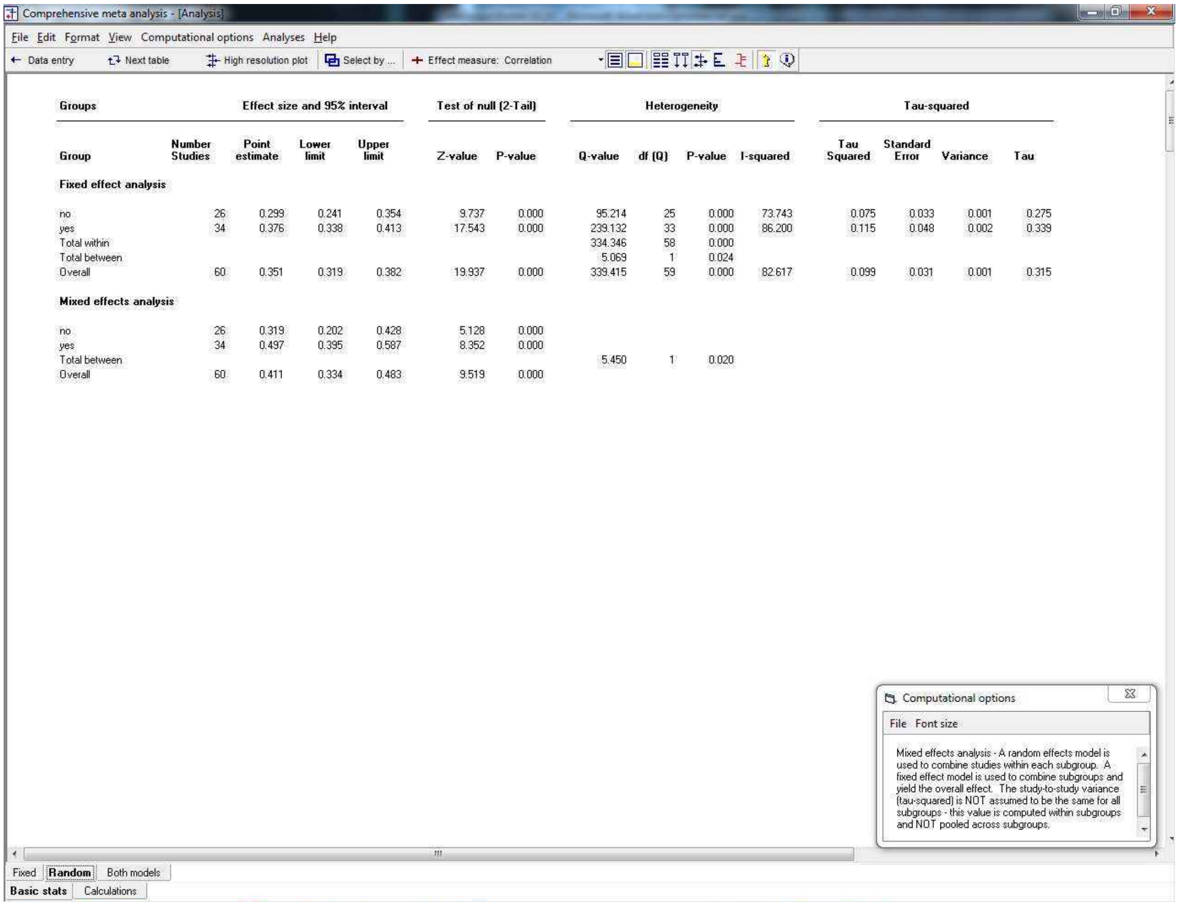
DOI: 10.1177/0956797614535810



Groups	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity				Tau-squared			
Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
Education	30	0.163	0.131	0.195	9.849	0.000	41.102	29	0.067	29.444	0.004	0.003	0.000	0.060
Games	6	0.483	0.418	0.543	12.715	0.000	5.920	5	0.314	15.542	0.002	0.008	0.000	0.045
Music	9	0.554	0.492	0.611	14.285	0.000	58.217	8	0.000	86.258	0.117	0.081	0.007	0.342
Sports	8	0.498	0.387	0.594	7.791	0.000	19.173	7	0.008	63.490	0.071	0.062	0.004	0.266
Total within							124.413	49	0.000					
Total between							159.516	3	0.000					
Overall	53	0.266	0.240	0.292	19.077	0.000	283.929	52	0.000	81.686	0.049	0.016	0.000	0.222
Mixed effects analysis														
Education	30	0.171	0.130	0.212	7.971	0.000								
Games	6	0.478	0.405	0.546	11.173	0.000								
Music	9	0.483	0.265	0.654	4.047	0.000								
Sports	8	0.533	0.344	0.680	4.944	0.000								
Total between							59.205	3	0.000					
Overall	53	0.245	0.209	0.280	13.038	0.000								

Figure S15. Screen Shot of the Results with Domain as a Moderator Variable (Additional Model 2: Excluding Team Sports, Only Solitary Deliberate Practice).

DOI: 10.1177/0956797614535810



Groups		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis														
no	26	0.299	0.241	0.354	9.737	0.000	95.214	25	0.000	73.743	0.075	0.033	0.001	0.275
yes	34	0.376	0.338	0.413	17.543	0.000	239.132	33	0.000	86.200	0.115	0.048	0.002	0.339
Total within							334.346	58	0.000					
Total between							5.069	1	0.024					
Overall	60	0.351	0.319	0.382	19.937	0.000	339.415	59	0.000	82.617	0.099	0.031	0.001	0.315
Mixed effects analysis														
no	26	0.319	0.202	0.428	5.128	0.000								
yes	34	0.497	0.395	0.587	8.352	0.000								
Total between							5.450	1	0.020					
Overall	60	0.411	0.334	0.483	9.519	0.000								

Figure S16. Screen Shot of the Results with Publication Status (i.e., Published vs. Unpublished) as a Moderator Variable (Sports Only)

References

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version 2*. Englewood, NJ: Biostat.

Cheung, S. F., & Chan, D. K-S. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology, 89,* 780–791.

Cheung, S. F. & Chan, D. K-S. (2008). Dependent correlations in meta-analysis: The case of heterogeneous dependence. *Educational and Psychological Measurement, 68*(5), 760–777.

Duval, S. J., & Tweedie, R. L. (2000a). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56(2),* 455–463.

Duval, S. J., & Tweedie, R. L. (2000b). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95(449),* 89–98.

Ericsson, K. A. (2013). 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.

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence, 45*, 34–45. doi:10.1016/j.intell.2013.04.001

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*, 2nd ed. Beverly Hills, CA: Sage.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

DOI: 10.1177/0956797614535810

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1*, 199–223.

Schmidt, F. L., Hunter, J. E., & Raju, N. S. (1988). Validity generalization and situational specificity: A second look at the 75% rule and Fisher's z transformation. *Journal of Applied Psychology*, *73*, 665–672.

Tuffiash, M., Roring, R. W., & Ericsson, K. A. (2007). Expert performance in SCRABBLE: Implications for the study of the structure and acquisition of complex skills. *Journal of Experimental Psychology: Applied, 13*(3), 124–134.

Young, B. W. (1998). Deliberate practice and the acquisition of expert performance in Canadian middle-distance running. Unpublished dissertation.

DOI: 10.1177/0956797614535810

OPEN PRACTICES DISCLOSURE

PLEASE COMPLETE AND RETURN TO EDITORIALOFFICE@PSYCHOLOGICALSCIENCE.ORG

***Psychological Science* manuscript #:** PSCI-13-1558.R2 **Corresponding author:** Brooke N. Macnamara

Articles accepted to *Psychological Science* after January 1, 2014, are eligible to earn badges that recognize open scientific practices: publicly available data, material, or preregistered research plans. Please read more about the badges here (https://osf.io/tvyxz/wiki/view/) and in the FAQs here (https://osf.io/tvyxz/wiki/faq/).

☐ **Please check this box if you are not interested in participating.**

To apply for one or more badges acknowledging open practices, please provide the information requested in the relevant sections below. **Your responses will be posted with your article as supplemental online material.**

You will not qualify for a badge for a given item unless you can provide a URL, doi, or other permanent path for accessing the specified information in a public, open-access repository AND answer "yes" to the relevant question #2. (Please see additional qualifiers in the section on preregistration.) If you answer "no" to question #2, but would still like your article to include a URL/doi/etc. to data/materials/registered research plans that are publicly available, we will publish the information in a note, but your article will not receive the associated badge.

Qualifying public, open-access repositories are committed to preserving data, materials, and/or registered analysis plans and keeping them publicly accessible via the web into perpetuity. Examples include the Open Science Framework (OSF; https://osf.io/) and the various Dataverse networks. Hundreds of other qualifying data/materials repositories are listed at http://re3data.org/ and http://databib.org/. Preregistration of an analysis plan must take place via a publicly accessible registry system (e.g., OSF, ClinicalTrials.gov or other trial registries in the WHO Registry Network, institutional registration systems). **Personal websites and most departmental websites do not qualify as repositories.**

There are, of course, circumstances in which it is not possible or advisable to share data, materials, or a research plan publicly. For example, there are cases in which sharing participant data could violate confidentiality. If you would like your article to include a note with an explanation of such circumstances, please provide the explanation in the Alternative Note section below.

Open Data:

1. Provide the URL, doi, or other permanent path for accessing the data in a public, open-access repository:

   osf.io/rhfsk

2. Is there sufficient information for an independent researcher to reproduce the reported results? If no, explain.

   yes

Open Materials:

1. Provide the URL, doi, or other permanent path for accessing the materials in a public, open-access repository:

2. Is there sufficient information for an independent researcher to reproduce the reported methodology? If no, explain.

DOI: 10.1177/0956797614535810

Preregistration:

1. Provide the URL, doi, or other permanent path to the registration in a public, open-access repository.*

2. Was the analysis plan registered prior to examination of the data or observing the outcomes? If no, explain.**

3. Were there additional registrations for the study other than the one reported? If yes, provide links and explain.*

4. Were there any changes to the preregistered analysis plan for the primary confirmatory analysis? If yes, explain.**

5. Are all of the analyses described in the registered plan reported in the article? If no, explain.*

*No badge will be awarded if (1) is not provided, **or** if (3) is answered "yes" without strong justification, **or** if (5) is answered "no" without strong justification.

**If the answer to (2) is "yes," the notation DE (Data Exist) will be added to the badge, indicating that registration postdates realization of the outcomes but predates analysis. If the answer to (4) is "yes" with strong justification for changes, the notation TC (Transparent Changes) will be added to the badge, indicating that the analysis plan was altered but the preregistered analyses and rationale for the change are provided.

Alternative Note: