

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/330400557>

Domain-General Models of Expertise: The Role of Cognitive Ability

Chapter · January 2019

DOI: 10.1093/oxfordhb/9780198795872.013.3

CITATIONS

0

READS

69

3 authors:



Zach Hambrick

Michigan State University

85 PUBLICATIONS 6,818 CITATIONS

[SEE PROFILE](#)



Alex Burgoyne

Michigan State University

10 PUBLICATIONS 62 CITATIONS

[SEE PROFILE](#)



Frederick L Oswald

Rice University

128 PUBLICATIONS 4,852 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Psychology of expertise [View project](#)



Expertise in chess [View project](#)

Oxford Handbooks Online

Domain-General Models of Expertise: The Role of Cognitive Ability

David Z. Hambrick, Alexander P. Burgoyne, and Frederick L. Oswald

The Oxford Handbook of Expertise

Edited by Paul Ward, Jan Maarten Schraagen, Julie Gore, and Emilie Roth

Subject: Psychology, Organizational Psychology, Cognitive Psychology

Online Publication Date: Jan 2019 DOI: 10.1093/oxfordhob/9780198795872.013.3

Abstract and Keywords

This chapter reviews evidence concerning the contribution of cognitive ability to individual differences in expertise. The review covers research in traditional domains for expertise research such as music, sports, and chess, as well as research from industrial-organizational psychology on job performance. The specific question that we seek to address is whether domain-general measures of cognitive ability (e.g., IQ, working memory capacity, executive functioning, processing speed) predict individual differences in domain-relevant performance, especially beyond beginning levels of skill. Evidence from the expertise literature relevant to this question is difficult to interpret, due to small sample sizes, restriction of range, and other methodological limitations. By contrast, there is a wealth of consistent evidence that cognitive ability is a practically important and statistically significant predictor of job performance, even after extensive job experience. The chapter discusses ways that cognitive ability measures might be used in efforts to accelerate the acquisition of expertise.

Keywords: expertise, skill acquisition, job performance, cognitive ability, intelligence

Domain-General Models of Expertise: The Role of Cognitive Ability

Why do some people reach higher levels of expertise in complex real-world tasks than other people? There is no doubt that domain-specific knowledge and skills contribute substantially to individual differences in expertise, whether it be in vocational or avocational pursuits (see Ward, Belling, Petushek, & Ehrlinger, 2017, for a review). Here,

while not denying the major importance of domain-specific factors, we consider the contribution of *domain-general cognitive ability factors*, reflecting the efficiency and effectiveness of basic mental processes.

Scope and Organization

In everyday life, people often rely on reputation to identify individuals with expertise—physicians, carpenters, auto mechanics, and so on. However, reputation does not ensure a high level of expertise (Ericsson, 2006). As a scientific concept, expertise is better thought of as a person's objective level of performance in a domain, as quantified by domain-relevant tasks (Ericsson & Smith, 1991) or proxy measures (e.g., performance-based rankings). For some domains, a single type of task may be sufficient to measure expertise. For example, given that playing good chess obviously depends on making good chess moves, chess expertise can be measured with *move-choice* tasks (de Groot, 1965/1978). For other domains, no single type of task captures expertise. For example, musical expertise comprises playing music from memory, sight-reading, and improvising, among other activities. Some musicians may be strong in all these activities; others may be strong in some but weak in others. Similarly, some auto mechanics may specialize in repairing diesel engines, others in transmissions, and still others in body repair. In short, expertise may be multidimensional.

Here, we review evidence for the role of cognitive ability in acquiring expertise. Along with limited space, there are two major reasons for this restricted focus. First, much of the controversy in contemporary research on expertise revolves around the question of whether, and to what extent, cognitive ability plays an important role in acquiring expertise (see, e.g., Detterman, 2014). Second, as industrial-organizational psychologists have demonstrated, measures of cognitive ability (along with other measures) are useful in organizational settings for selecting job applicants, because they are consistently and positively correlated with job performance (Schmidt & Hunter, 1998, 2004). Similarly, scores on standardized cognitive tests such as the Graduate Record Examination (GRE), the Law School Admission Test (LSAT), and the Graduate Management Admission Test (GMAT) are useful and valid predictors of success in advanced academic studies (Kuncel & Hezlett, 2007).

Table 1 Domain-general cognitive ability factors, with representative definitions and examples of assessments

Construct	Definition/tests
Intelligence	Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think

Domain-General Models of Expertise: The Role of Cognitive Ability

	abstractly, comprehend complex ideas, learn quickly, and learn from experience. (Gottfredson, 1997, p. 13)
	- Wechsler Adult Intelligence Scale (full-scale IQ)
	- Raven's Progressive Matrices (fluid intelligence, or Gf)
	- Air Force Officer Qualifying Test (AFQT)
Executive functioning	Executive function can be thought of as the set of abilities required to effortfully guide behavior toward a goal, especially in nonroutine situations. (Banich, 2009, p. 89)
	- Wisconsin Card Sorting task
	- Tower of Hanoi
	- Trailmaking
Working memory capacity	[Working memory capacity refers to] the attentional processes that allow for goal-directed behavior by maintaining relevant information in an active, easily accessible state outside of conscious focus, or to retrieve that information from inactive memory, under conditions of interference, distraction, or conflict. (Kane et al., 2007, p. 23)
	- Operation span
	- n-back
	- Backward digit span
Attentional control	Attention control refers to the ability to protect items that are actively being maintained in working memory, to effectively select target representations for active maintenance, and to filter out irrelevant distractors and prevent them from gaining access to working memory. (Unsworth, Fukuda, Awh, Vogel, 2015, p. 864)
	- Attention Network Task (ANT)
	- Stroop task
	- Flanker task

Speed of processing	Processing speed refers to the ability to quickly and efficiently carry out mental operations. (Tucker-Drob, 2011, p. 333)
	- Digit-symbol substitution
	- Letter/pattern comparison
	- Choice reaction time

Table 1 lists the cognitive ability constructs that we consider, with a definition of each construct and examples of assessments. Though often treated as if they are empirically and conceptually distinct, measures of these constructs correlate positively, and sometimes near 1.0 after correcting for measurement error (e.g., Engelhardt et al., 2016; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). This implies that common mechanisms underlie individual differences in these constructs, which could include acquired factors such as general problem-solving strategies, neural factors such as the functional connectivity of different brain regions, and genetic factors (see Haier, 2016). Any (or all) of these factors could contribute, directly or indirectly, to associations between cognitive ability factors and expertise. Cognitive ability constructs are also sometimes described as being *innate*, but heritability (i.e., estimated genetic contribution) of any human characteristic is always less than 100 percent (Turkheimer, 2000), leaving room for a contribution of environmental factors. At the population level, heritability is typically around 50 percent for measures of cognitive ability, indicating roughly equal contributions of genetic and environmental factors to individual differences (Knopik, Neiderhiser, DeFries, & Plomin, 2016).

Review of Evidence for Role of Cognitive Ability in Expertise

Classical theories of skill acquisition (e.g., Fitts & Posner, 1967) posit that domain-general processes impact performance early in training, after which procedural knowledge becomes the major determinant. Consistent with this assumption, there is ample evidence that cognitive ability predicts initial acquisition of knowledge/skill in complex domains. For example, measures of cognitive ability from test batteries such as the Armed Services Vocational Aptitude Battery (ASVAB) positively predict job training performance, with validity coefficients averaging around 0.50 (Schmidt & Hunter, 2004). It is less clear whether cognitive ability remains a valid predictor of performance differences after extensive training. This question is not only of theoretical interest to expertise researchers (e.g., the *circumvention-of-limits hypothesis*; Hambrick & Meinz, 2011), but one of applied interest: If a measure significantly predicts performance in a task,

especially beyond the beginner level, then that measure might be used to help make decisions such as whom to select for a costly training program.

Next, we review evidence relevant to this question. We performed literature searches in Google Scholar and PsycINFO, using a wide range of search terms (e.g., “expertise” and “cognitive ability” with “sports,” “chess,” and “aviation”). We searched approximately 1,300 documents, identifying relevant studies in two primary literatures: the literature on expertise in domains such as music, chess, and sports, and the literature on job performance. Our review focuses on studies that tested cognitive ability–performance relations across different levels of expertise, or at least in non-beginners. (We excluded studies that measured specific aptitudes, such as music aptitude and mechanical aptitude.) The specific question we set out to address is whether expertise mitigates the effect of cognitive ability on domain-relevant performance. Throughout the chapter, we note correlations between domain-specific factors and domain-relevant performance for comparative purposes.

Games

There is evidence that cognitive ability predicts acquisition of chess skill at the beginner level (e.g., de Bruin, Kok, Leppink, & Camp, 2014), but it is unclear what role it plays at higher levels of expertise. Evidence is mixed. For example, in two studies, Unterrainer and colleagues (Unterrainer, Kaller, Halsband, & Rahm, 2006; Unterrainer, Kaller, Leonhart, & Rahm, 2011) found near-zero correlations between IQ measures and chess rating in small samples of chess players ($N = 25$ and 26 , respectively) with intermediate-level average chess ratings, whereas Grabner, Stern, and Neubauer (2007) found a correlation of 0.35 between IQ and chess rating using a larger sample ($N = 90$) with a slightly higher average rating. Even the latter finding is tentative because the confidence interval around a correlation of 0.35 with a sample of 90 is quite wide, ranging from 0.15 to 0.52 .

To try to make sense of the conflicting evidence, Burgoyne et al. (2016) performed a meta-analysis of the relationship between cognitive ability and chess expertise. Across 19 studies, four cognitive abilities were measured: fluid intelligence, crystallized intelligence, short-term/working memory, and processing speed. The meta-analytic average of the correlations was 0.22 ($p < 0.001$). (Correlations between chess rating and domain-specific factors are typically much larger (e.g., Pfau & Murphy, 1988, $r = 0.68$).) Burgoyne et al. also found that the correlation between fluid intelligence and expertise was stronger for less skilled (unranked) chess players than for more skilled (ranked) players (0.33 vs 0.10 ; see Burgoyne et al., in press, for a correction to the originally reported values). However, it is important to note that expertise was highly confounded both with age (i.e., nearly all ranked chess players were adults, nearly all unranked chess players were youths) and with type of skill measure (i.e., Elo ratings for ranked players, chess test scores for unranked players).

In another meta-analysis, Sala et al. (2017) found that chess players are, on average, significantly higher in measured cognitive ability than non-chess players. As most of the chess samples included relatively highly skilled players, this could be because people high in cognitive ability are more likely to enjoy success in chess than those lower in cognitive ability, and are thus less likely to quit the game (i.e., performance effects). Alternatively, it could be that playing chess enhances cognitive ability (i.e., training effects) or because higher-ability individuals are more likely to take up chess than lower-ability individuals (i.e., selection effects).

Summarized, evidence is inconclusive on whether the importance of cognitive ability declines with chess expertise. The same is true for other games. In neuroimaging studies, Lee et al. (2010, $N = 16$) and Jung et al. (2013, $N = 17$) reported IQ data on small samples of elite Baduk (Korean for Go) players. Full-scale IQ was lower for the Baduk players than for a control group by 8 points in Lee et al. ($M = 93.2$ vs 101.2 , $p = 0.052$) and 7.7 points in Jung et al. ($M = 93.1$ vs 100.8 , $p = 0.06$). The fact that the Baduk group in each study had a *lower* average IQ than the control group is somewhat puzzling and may partly reflect the fact that the Baduk group had less education on average than the control group (by 1.3 years in Lee et al., $p < 0.05$; and by 1.1 years in Jung et al., $p = 0.19$).

A much larger study of Go was carried out by Masunaga and Horn (2001). Participants ($N = 263$) representing wide ranges of Go expertise completed tests of both domain-general and domain-specific factors. The domain-general battery included standard tests of fluid reasoning, short-term memory, and perceptual speed; the domain-specific battery included *Go-embedded* tests designed to measure the same abilities but with Go-specific content. The Go reasoning test was modeled after move-choice tasks in chess (de Groot, 1965/1978), and can be considered a measure of Go skill. On average, the domain-general measures correlated 0.18 with Go move-choice. The highest correlations were for fluid intelligence (avg. $r = 0.27$); group average r values (obtained from Takagi, 1997) were as follows: beginner (avg. $r = 0.21$, $p = 0.001$, $n = 62$), intermediate (avg. $r = 0.33$, $p < 0.001$, $n = 89$), expert (avg. $r = 0.27$, $p < 0.001$, $n = 92$), and professional (avg. $r = 0.18$, $p = 0.14$, $n = 20$). These correlations are not significantly different from each other (z statistics < 1). The average correlations between the fluid intelligence measures and Go rank were non-significant: beginner (avg. $r = -0.03$), intermediate (avg. $r = -0.06$), expert (avg. $r = 0.03$), and professional (avg. $r = -0.26$). It is somewhat surprising that fluid intelligence correlated with move-choice performance but not with Go rank, given the high correlation between the latter measures ($r = 0.71$) and that move-choice must be critical for success in Go tournaments. It could be that the move-choice task was somewhat artificial in that it presented the player with novel positions, whereas in actual Go games a skilled player can steer a game toward familiar territory and thus encounter more familiar positions. The average correlation of the domain-specific measures with Go move-choice was 0.46 and with Go rank was 0.47.

Word games have also been used to investigate the relationship between cognitive ability and expertise. Tuffiash, Roring, and Ericsson (2007) compared groups of elite, average, and novice Scrabble players on tests of various domain-specific and domain-general

cognitive abilities. There were significant group differences (favoring higher expertise) in the domain-relevant tasks (e.g., anagramming; medium-to-large effect sizes), but not in domain-general perceptual speed (i.e., digit-symbol substitution). However, the rated players (average and elite groups) outperformed the novices on tests of vocabulary and reading comprehension ($d_s > 2$). More recently, Toma, Halpern, and Berger (2014) found that Scrabble and crossword puzzle experts tended to outperform the control subjects on two tests of working memory capacity (avg. $d = 1.23$). As in chess, these skill group differences could reflect performance effects, training effects, and/or selection effects.

There have been a few studies of poker expertise. In a study of undergraduate students described as being *familiar with Texas Hold 'em poker*, Leonard and Williams (2015) found that scores on several subtests from the Stanford-Binet Intelligence Scales correlated non-significantly with performance on a test of poker skills. However, in a sample of 155 undergraduates representing a wider range of Texas Hold 'em experience, Mainz et al. (2012) found that working memory capacity explained a significant amount of variance (avg. $R^2 = 0.071$) in measures of Hold 'em component skills (e.g., hand evaluation), above and beyond poker knowledge (avg. $R^2 = 0.358$). Moreover, there was no evidence for Poker Knowledge \times Working Memory Capacity interactions, indicating that effects of working memory capacity on performance were similar across levels of poker knowledge.

Finally, Ceci and Liker (1986) found that groups of *nonexperts* ($n = 16$) and *experts* ($n = 14$) in horserace handicapping were not only nearly identical in average IQ, but both near the population mean of 100 ($M_s = 99.3$ and 100.8 , respectively). However, in a re-analysis, Detterman and Spry (1988) found that the correlation between IQ and a key measure of success (correct top horse) was positive in the expert group ($r = 0.35$, or 0.59 after correction for unreliability) but negative in the novice group ($r = -0.25$, or -0.42 after correction for unreliability), casting some doubt on the argument that IQ is unrelated to success in horserace handicapping. That said, these sample sizes were very small, and the result would obviously need to be replicated in a larger sample.

Music

It is also unclear what role cognitive ability plays in music expertise beyond the beginner level. Ruthsatz, Detterman, Griscom, and Cirullo (2008) found that scores on a test of fluid intelligence (Raven's Progressive Matrices) correlated positively and significantly with musical achievement in high school band members ($r = 0.25$, $n = 178$), but not in university music majors ($r = 0.24$, $n = 19$) and music institute students ($r = 0.12$, $n = 64$)—although statistical power obviously differed across the samples. Moreover, the correlations did not differ between the lower- and higher-skill groups (tests of differences in r_s , z statistics < 1). Correlations with estimated amount of *deliberate practice* (Ericsson, Krampe, & Tesch-Römer, 1993) in the high school, university, and music institute samples were 0.34 , 0.54 , and 0.31 , respectively (all significant).

Meinz and Hambrick (2010) had pianists provide estimates of deliberate practice and perform tests of both working memory capacity and sight-reading. Deliberate practice accounted for 45 percent of the variance in sight-reading performance; working memory capacity accounted for an additional 7.4 percent. (The correlation between deliberate practice and working memory capacity was near zero.) Moreover, the Deliberate Practice \times Working Memory Capacity interaction was non-significant, indicating that the effect of working memory capacity on performance was similar across levels of deliberate practice. By contrast, perceptual speed did not contribute significantly to the prediction of sight-reading performance.

Using a sample of 52 pianists with a more uniform level of skill (piano majors at a music university), Kopiez and Lee (2008) found that although correlations between sight-reading performance and fluid reasoning ($r = 0.12$) and reaction time (avg. $r = -0.07$) were non-significant, there was a significant correlation for a measure of perceptual speed ($r = -0.44$; faster processing, higher performance). The correlation between working memory and sight-reading performance did not reach significance ($r = 0.26$, $p = 0.062$). Correlations between measures of domain-relevant motoric speed (*trilling*) and sight-reading performance averaged 0.50; the correlation between deliberate practice and sight-reading performance was 0.50.

Other studies have compared musicians of varying levels of skill on measures of cognitive ability, as well as musicians to non-musicians. Schellenberg and colleagues have found that musically trained individuals tend to be higher in full-scale IQ than non-musically trained individuals (see Schellenberg & Weiss, 2013, for a review). As with chess, this difference could reflect selection effects, training effects, and/or performance effects.

Sports

Evidence for the role of cognitive ability in sports expertise is inconsistent, as well. For example, Lyons, Hoffman, and Michel (2009) found that scores on the Wonderlic IQ test correlated near zero ($r = -0.04$) with future NFL performance in a large sample of elite college football players (total $N = 762$; see also Berri & Simmons, 2011), whereas Vestberg, Gustafson, Maurex, Ingvar, and Petrovic (2012) found that a measure of executive functioning (design fluency from the D-KEFs) significantly predicted goals scored in elite Swedish soccer players ($r = 0.54$, $N = 25$), albeit in a much smaller sample.

In a meta-analysis of 42 studies, Mann, Williams, Ward, and Janelle (2007) compared nonexpert and expert athletes on performance measures from sports-specific perceptual-cognitive tasks (e.g., occlusion paradigms). Across measures, there was a statistically significant advantage for experts ($ds = 0.23$ to 0.35). Given evidence for the importance of training in acquiring skill in sports (e.g., Ward, Hodges, Starkes, & Williams, 2007), these differences likely reflect domain-specific factors, but they could also reflect domain-general factors as well (Ward et al., 2017).

In a subsequent meta-analysis of 20 studies, Voss, Kramer, Basak, Prakash, and Roberts (2010) found a significant advantage for athletes over non-athletes on processing speed (Hedge's $g = 0.67$) and *varied* attention tasks (Hedge's $g = 0.53$) but not attentional cueing (Hedge's $g = 0.17$). (Hedge's g is similar to Cohen's d .) These results lend some support to the possibility that domain-general factors contribute to sports expertise (i.e., performance effects), but as before could also reflect selection effects and/or training effects (Ward et al., 2017).

Science

The relationship between cognitive ability and scientific expertise has also been of interest to psychologists. Early studies of this relationship yielded mixed evidence. Bayer and Folger (1966) reported a correlation of -0.05 between IQ and number of citations (a proxy for scientific expertise) in a sample of 224 biochemists, and Folger, Astin, and Bayer (1970) found correlations ranging from 0.04 to 0.10 between cognitive ability in high school and number of citations in a sample of 6,300 PhDs. However, Creager and Harmon (1966) found that scores on the GRE predicted citation counts 8–12 years later (median $r = 0.28$; cited in Clark & Centra, 1982) in NSF predoctoral fellowship applicants (see also Kaufman, 1972).

More convincing results come from a meta-analysis of 6,589 correlations from 1,753 independent samples by Kuncel, Hezlett, and Ones (2001). After applying psychometric corrections for statistical artifacts such as range restriction and measurement unreliability in the criterion measures, Kuncel et al. found that estimated validity coefficients (ρ s) in the population for the General GRE test were positive and significant not only for first-year GPA (avg. $\rho = 0.36$; avg. $r = 0.24$) and overall GPA (avg. $\rho = 0.34$; avg. $r = 0.23$), but also for publication citation counts (avg. $\rho = 0.20$; avg. $r = 0.15$), and were positive for research productivity (avg. $\rho = 0.10$; avg. $r = 0.08$). Validities for the Subject GRE test (reflecting domain-specific knowledge) were higher for all outcomes, including publication citation counts ($\rho = 0.24$; $r = 0.20$) and research productivity ($\rho = 0.21$; $r = 0.17$).

This evidence corroborates the results of the Study of Mathematically Precocious Youth (SMPY). As part of a planned 50-year study, the Scholastic Aptitude Test (now just called the SAT) was administered to a large national sample of gifted youth by age 13, and those scoring in the top 1 percent were tracked into adulthood ($N > 2,300$). Analyses have since demonstrated that—even within this highly restricted range of ability—SAT scores are positively predictive of success in scientific fields. For example, Lubinski (2009) found that, compared with individuals in the 99.1 percentile, those in the 99.9 percentile were about 5 times more likely to have published in a STEM journal and about 3 times more likely to have been awarded a patent.

Thus, there is evidence that cognitive ability predicts general measures of scientific expertise. There is, however, some evidence that cognitive ability may become less important in specific scientific tasks. Hambrick et al. (2012) had a sample of 67 participants representing a wide range of knowledge and experience in geological fields perform a highly realistic *bedrock mapping* task in which the goal was to create a *field map* representing the geological structure of an area based on observable features (e.g., rock outcrops). There was a significant Geological Knowledge \times Visuospatial Ability interaction, such that a composite measure of visuospatial ability positively predicted map accuracy, but only in those with lower levels of geological knowledge.

Surgery/Medicine

There is a growing literature on the role of cognitive ability in surgical expertise, but the results are no clearer than in other domains. In a study of 120 surgical residents (Schueneman, Pickleman, Hesslein, & Freeark, 1984), 4 of 5 measures of visuospatial ability correlated significantly with surgical performance (avg. $r = 0.28$), as evaluated by attending surgeons. Year of residency correlated 0.60 with surgical performance. Gibbons, Baker, and Skinner (1986) found that scores on a hidden figures test correlated significantly with surgical performance in small samples of surgical residents ($r_s = 0.55$ and 0.60 , $N_s = 42$ and 16), but Deary, Graham, and Maran (1992) found no significant positive correlations between expert ratings of surgical ability and intelligence test scores in trainee surgeons ($N = 22$).

Several studies have compared ability–performance correlations across different levels of surgical expertise. Wanzel et al. (2003) found that scores on two tests of “high-level” visuospatial ability (mental rotation and surface development) correlated significantly with expert ratings of surgical performance in dental students (*novices*, $n = 27$, avg. $r = 0.56$), but not in surgical residents (*intermediates*, $n = 12$) or staff surgeons (*experts*, $n = 8$). The correlations for the latter groups were not reported, but given the extremely small sample sizes here, they would not be significantly different from the novice correlation even if they were assumed to be zero. Comparing groups of surgeons on a simulated videoscopic task, Keehner et al. (2004) found that a measure of visuospatial ability correlated significantly with mean skill rating in a low experience group ($r = 0.39$, $n = 48$), but not in a high experience group ($r = 0.02$, $n = 45$). But, again, the correlations are not significantly different ($z = 1.83$, $p = 0.067$).

Gallagher, Cowie, Crothers, Jordan-Black, and Satava (2003) found that scores on a test of visuospatial ability in which participants recover three-dimensional structures from two-dimensional images correlated significantly and similarly with performance on a laparoscopic laboratory cutting task in two samples of novices ($r_s = 0.50$ and 0.50 , $n_s = 48$ and 32) and in experienced surgeons ($r = 0.54$, $n = 18$). These correlations also do not differ across skill level. Enochsson et al. (2006) compared 18 resident and 11 expert surgeons in a simulated gastroscopy task, and found that correlations between scores on

a test of visuospatial ability (card rotation test) and various metrics of performance in these very small samples were generally non-significant for both groups (avg. $r = 0.06$).

Murdoch, Bainbridge, Fisher, and Webster (1994) found that both manual dexterity and visuospatial ability correlated significantly with medical students' performance on microsurgical tasks ($r_s = -0.54$ and 0.36 , respectively). And in a sample of surgeons ($N = 94$), Risucci, Geiss, Gellman, Pinard, and Rosser (2001) found that measures of visuospatial ability correlated moderately (and 10/12 significantly) with performance on four surgical tasks (avg. $r = -0.30$; higher ability, faster performance); a measure of domain-specific experience correlated significantly with two of the performance measures ($r_s = 0.35$ and 0.29), as did a measure of domain-specific knowledge (post-test examination; $r_s = 0.30$ and 0.39). Groenier, Schraagen, Miedema, and Broeders (2014) examined the validity of tests of cognitive ability for predicting performance in a laparoscopic training simulator in medical students ($N = 53$) over 2 months. In univariate analyses, visuospatial ability, spatial memory, perceptual speed, and reasoning ability significantly predicted one performance measure (motion efficiency), while visuospatial ability and reasoning ability predicted another performance measure (duration). By contrast, in multivariate analyses, which controlled for correlations among the predictor variables, only one of the preceding effects was significant. The finding that univariate effects became non-significant in the multivariate analyses suggests that variance common to the ability measures (a g factor) may have been predictive of surgical performance.

More recently, Louridas and colleagues performed a meta-analysis of 52 studies on the relationship between various measures of cognitive ability and performance in laparoscopic, open, and endoscopic surgery (Louridas, Szasz, de Montbrun, Harris, & Grantcharov, 2016). Only a few cognitive ability measures positively predicted surgical performance across multiple studies, among them the mental rotation test, a pictorial surface orientation test, and the grooved pegboard test. Louridas et al. concluded that "no single test has been reported to reliably predict technical performance across the range of techniques and skills required of surgical trainees" (p. 689).

One other study fits in this category. In a sample ($N = 428$) that included professionals in exercise science-related jobs (e.g., physicians, trainers) as well as participants from the general population, Petushek, Cokely, Ward, and Myer (2015) found that two measures of cognitive ability had non-significant effects on performance in a task designed to assess risk of injury to the anterior cruciate ligament (ACL). By contrast, domain-specific factors (i.e., ACL knowledge and use of particular visual cues) were positive and statistically significant predictors of performance ($r = 0.59$ for ACL knowledge; Petushek, 2014).

Aviation

Several studies have tested for cognitive ability–performance correlations in aviation. In a sample of 86 pilots representing a wide range of experience and skill, along with 96 non-pilots, Morrow, Menard, Stine-Morrow, Teller, and Bryant (2001) found that a cognitive ability composite (working memory, perceptual speed, and visuospatial ability) positively predicted aviation-related performance (i.e., a composite reflecting accuracy in recalling and understanding air traffic control, ATC, commands), accounting for 29 percent of the variance. An expertise composite (ATC knowledge and flight hours) accounted for an additional 37 percent of the variance, but the Expertise \times Cognitive Ability interaction was non-significant for all performance measures, indicating that the effect of cognitive ability on performance was similar across levels of expertise.

In a similar study of pilots ($N = 91$), Morrow et al. (2003) found that a cognitive ability composite accounted for an average of 22 percent of the variance in ATC tasks; an expertise composite accounted for an additional 28 percent of the variance, on average. (Expertise \times Cognitive Ability interactions are not reported for this study.) Consistent with these findings, in a study of 97 licensed pilots with a wide range of flight experience, Taylor, O'Hara, Mumenthaler, Rosen, and Yesavage (2005) found that performance in an aviation communication task correlated significantly with working memory ($r = 0.76$), processing speed ($r = 0.33$), and interference control ($r = 0.43$), but interactions of expertise (flight rating) with these factors were all non-significant.

Using a sample with 25 novice and 25 expert pilots, Sohn and Doane (2003) found that working memory capacity predicted success in an aviation situational awareness task, but only in pilots who scored low on an aviation-specific test measuring skilled access to long-term memory (i.e., long-term working memory; Ericsson & Kintsch, 1995), as evidenced by a significant Long-Term Working Memory \times Working Memory Capacity interaction. In a similar study, Sohn and Doane (2004) found that two measures of working memory capacity (spatial span and verbal span) correlated more strongly with situational awareness in 25 novice pilots ($r_s = 0.52$, $p < 0.01$, and 0.30 , respectively) than in 27 expert pilots ($r_s = 0.10$ and 0.10 , respectively). However, these correlations are not significantly different from each other across skill groups (z statistics < 1.7). Sohn and Doane (2004) did not test the Long-Term Working Memory \times Working Memory Capacity interaction using the full sample (as in their earlier study), but instead tested it separately in each skill group, finding significance only in the expert group.

Finally, in a small sample of private pilots ($N = 24$), Causse, Dehais, and Pastor (2011) examined the relationship of broad cognitive abilities (reasoning and processing speed) and executive functions (working memory updating, set-shifting, and inhibition) to performance during a 45-minute flight simulator task. Reasoning ability correlated significantly with flight-path deviations ($r_s = -0.63$); the other correlations were non-significant.

Job Performance

Measures of general cognitive ability positively predict job performance (Schmidt & Hunter, 2004), but do they remain valid predictors after extended job experience? This question has long been of interest to industrial-organizational psychologists. Using laboratory perceptual-motor tasks, Fleishman and colleagues demonstrated that general ability factors become less important with training, whereas task-specific factors become more important (e.g., Fleishman & Rich, 1963; see Hulin, Henry, & Noon, 1990, for other examples). However, the general finding from large-scale studies of actual work performance (as opposed to laboratory tasks) is that cognitive ability remains a significant predictor of job performance even after extensive job experience.

McDaniel (1986) investigated the impact of job experience on the validity of general cognitive ability using the General Aptitude Test Battery (GATB) database.¹ Compiled by the U.S. Employment Service in the 1970s, this database includes information on a large sample of civilian workers, including measures of job performance (i.e., supervisor ratings), job experience, and cognitive ability. McDaniel computed correlations between “intelligence” scores from the GATB (based on visuospatial, vocabulary, and arithmetic reasoning scores) and job performance across different levels of job experience. As shown in Figure 1, the correlations decrease somewhat as a function of job experience, but are still significant at the maximum amount of job experience (10+ years, $r = 0.20$, corrected $r = 0.29$).

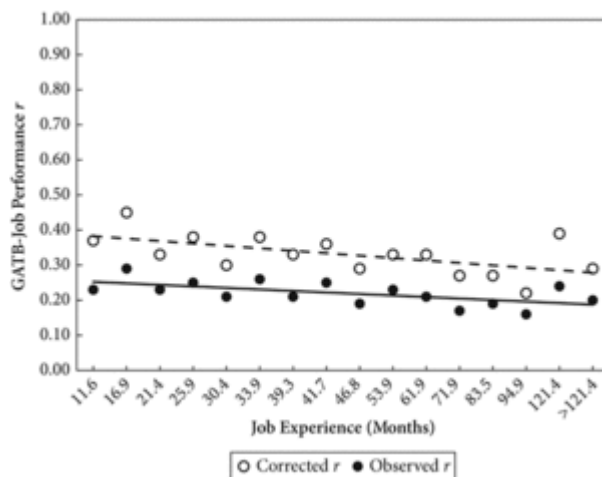


Figure 1. Correlations between GATB intelligence scores and job performance ratings as a function of job experience (total $N = 16,058$; across intervals, $n_s = 1,000$ to 1,050, except for >121.4 months, $n = 879$). Solid circles represent observed (raw) correlations; open circles represent correlations after correction.

Data from McDaniel, M. A., "The evaluation of a causal model of job performance: The interrelationships of general mental ability, job experience, and job performance," Tables 1 and 19, PhD thesis, George Washington University, Washington, D.C., 1986.

Again using the GATB database, Farrell and McDaniel (2001) extended Ackerman's (1988) model of skill acquisition to job performance. Briefly, Ackerman hypothesized that involvement of different cognitive abilities in skill acquisition is moderated by the consistency of the task: When the demands of the task are consistent, meaning that the stimuli, rules, and sequences of action remain constant, automaticity can develop and the influence of general cognitive ability (reflecting attentional resources) on performance decreases with training.

Meanwhile, the influence of perceptual speed increases and later decreases (i.e., an inverted U function) and the influence of psychomotor speed increases. To test this model, Farrell and McDaniel classified jobs as consistent or inconsistent using two different definitions of consistency: the complexity of the job (low complexity = consistent, high complexity = inconsistent) and tolerance for repetition required to perform the job (high tolerance for repetition = consistent, low tolerance for repetition = inconsistent). They then computed correlations between two cognitive composites (along with psychomotor speed) from the GATB (intelligence and perceptual speed) and job performance for different levels of job experience. Support for Ackerman's model was mixed. For example, the intelligence correlations decreased as a function of job experience for low complexity jobs, but *increased* slightly for high tolerance for repetition jobs. For the present discussion, the more important finding is simply that the cognitive ability factors significantly predicted job performance even at the maximum level of job experience: intelligence (avg. $r = 0.25$; avg. corrected $r = 0.34$) and perceptual speed (avg. $r = 0.15$; avg. corrected $r = 0.20$).

Studies of military personnel provide additional evidence that cognitive ability remains a significant predictor of job performance beyond initial training. Schmidt, Hunter, Outerbridge, and Goff (1988) tested for effects of cognitive ability and job experience on job performance in a sample of 1,474 soldiers in four jobs (armor repairman, armor crewman, supply specialist, and cook). Job performance was measured using work

samples and supervisor ratings; cognitive ability was measured using the Armed Forces Qualification Test (AFQT) score from the ASVAB, which is based on Arithmetic Reasoning, Mathematics Knowledge, Paragraph Comprehension, and Word Knowledge subtests. (Job knowledge was also treated as a measure of job performance, though we think of it as a *predictor* of job performance.) Up to 5 years of job experience, correlations between AFQT scores and job performance were nearly constant. Across this span, correlations ranged from 0.38 to 0.42 for work samples and from 0.18 to 0.36 for supervisor ratings. Beyond 5 years of job experience (i.e., 61+ months), there was apparent convergence of ability groups for most measures, indicating a drop in validity beginning at 5 years. However, average amount of job experience was actually much higher than 5 years in this group—from 9.5 years to 13 years, depending on the job. Moreover, only 1 of 12 AFQT × job experience interactions (work sample performance for armor crewman) was statistically significant, and it was not clearly interpretable as supporting convergence of the ability groups. Schmidt et al. concluded that “[a]t least out to 5 years, the validity of general mental-ability measures appears neither to decrease ... nor to increase Instead, the validity remains relatively constant” (p. 56).

Wigdor and Green (1991) reported results of the Joint-Service Job Performance Measurement/Enlistment (JPM) Standards Project, a large study initiated in 1980 by the U.S. Department of Defense to develop measures of military job performance. Wigdor and Green reported that, across 23 jobs ($N = 7,093$ military personnel), the median correlation between AFQT scores and *hands-on* job performance (HOJP) was 0.26 (0.38 after correction for range restriction). They also reported mean hands-on performance for four AFQT categories (representing different levels of cognitive ability) as a function of job experience. As shown in Figure 2, mean differences among AFQT categories were largest at 0–12 months (about 10 points, or 1 *SD*), but still sizeable thereafter (5–6 points, or 0.50–0.60 *SD*). Wigdor and Green concluded that “the level of performance is positively related to AFQT score category at each of the four levels of job experience” (p. 163) and noted that “the lowest aptitude group never reaches the initial performance level of the highest aptitude group” (p. 163).

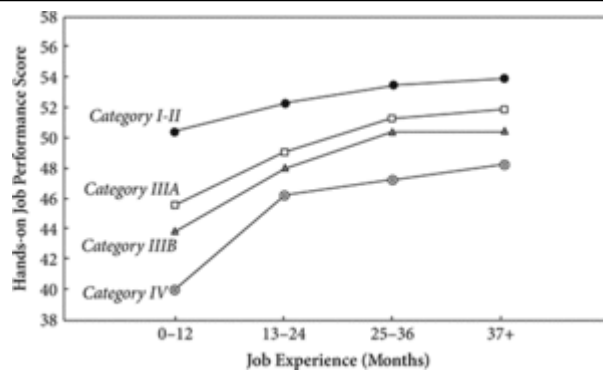


Figure 2. Mean Hands-on Job Performance Score by AFQT category (i.e., cognitive ability level). Percentile ranges for AFQT categories: I-II (65–99), IIIA (50–64), IIIB (31–49), and IV (10–30) (see Wigdor & Green, 1991, p. 53).

Data from Wigdor, Alexandra K., and Green, Bert F., *Performance assessment for the workplace*, Volume 1, p. 53, Table 2.5, National Academy Press, 1991.

To further investigate cognitive ability–job performance relations, we obtained the JPM dataset.² The final dataset included 31 jobs and a total sample size of 10,088 military personnel. We performed three new analyses. First, we computed the AFQT–HOJP correlation across the job experience intervals used by Wigdor and Green (1991). As shown in Figure 3A, the correlations are as follows: 0–12 months ($r = 0.34$, $p <$

0.001 , $n = 747$), 13–24 months ($r = 0.21$, $p < 0.001$, $n = 5,234$), 25–36 months ($r = 0.19$, $p < 0.001$, $n = 2,338$), and 37+ months ($r = 0.22$, $p < 0.001$, $n = 1,769$).³ There is a statistically significant drop in the correlation from the first year of service to the second ($z = 3.60$, $p < 0.001$), but AFQT is still a statistically significant predictor of individual differences in HOJP after the first year of service. Second, capitalizing on the large data set, we broke the 37+–month group into additional experience intervals, to 85+ months (creating any more groups would result in small sample sizes, $ns < 50$). As shown in Figure 3B, the AFQT–HOJP correlation decreases from the first year to the second, stabilizes, and then increases—though the estimates become less precise as sample size decreases.

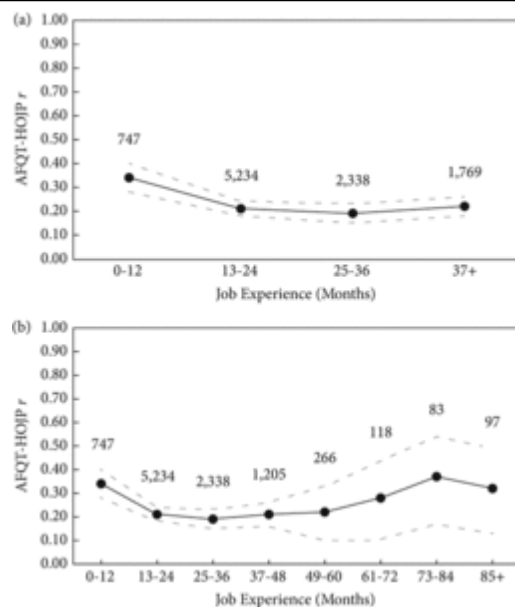


Figure 3. Correlations (with 95 percent confidence intervals) between AFQT scores and Hands-on Job Performance (HOJP) scores at 4 (A) and 8 (B) job experience intervals. Dashed lines are 95 percent confidence intervals; adjacent values are sample sizes.

Data from Joint-Service Job Performance Measurement/Enlistment (JPM) Standards Project ($N = 10,088$).

Finally, as the most statistically powerful analysis, we evaluated the Job Experience \times AFQT interaction on HOJP using the entire data set via moderated multiple regression. (Prior to performing the regression analysis, we log-transformed job experience because it was non-normal, skewness = 2.40 and kurtosis = 9.56, and we mean-centered the predictors.) There were significant main effects of both AFQT ($\beta = 0.210$, $t = 21.92$, $p < 0.001$, part $r^2 = 0.044$) and log job experience ($\beta = 0.167$, $t = 17.37$, $p < 0.001$, part $r^2 = 0.028$) on HOJP. High levels of both AFQT and

job experience were associated with higher HOJP. The AFQT \times Log Job Experience interaction was also statistically significant and under-additive ($\beta = -0.023$, $t = -2.41$, $p = 0.016$, part $r^2 = 0.0005$), though the effect was virtually nil, indicating that AFQT was predictive of HOJP regardless of level of job experience (see Figure 4).

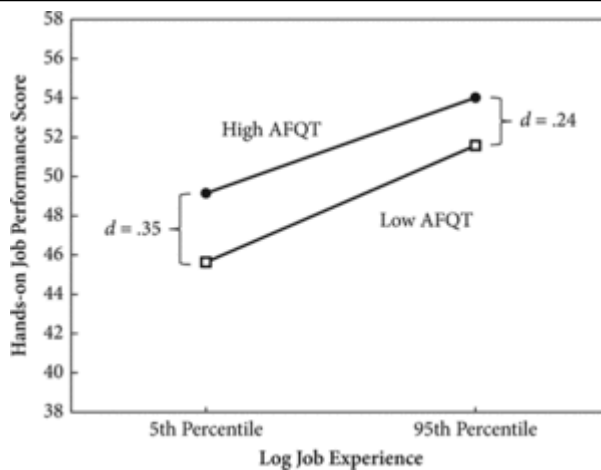


Figure 4. Predicted values for Hands-on Job Performance (HOJP) for low vs high AFQT (25th vs 75th percentile) at 5th vs 95th percentiles for Log Job Experience. Values generated using regression equation with AFQT score (mean-centered), Log Job Experience (mean-centered), and AFQT \times Log Job Experience as predictors of HOJP score: $\text{HOJP score} = 50.031 + 0.096(\text{AFQT}) + 8.119(\text{Log Job Experience}) + -0.052(\text{AFQT} \times \text{Log Job Experience})$. $N = 10,088$.

The overall picture to emerge from these large-scale studies is that cognitive ability remains a significant predictor of job performance, even after extensive job experience and even if validity drops initially. The question of how *far* beyond initial training cognitive ability predicts job performance is unanswered, but the results we have just reviewed indicate at least 5 years (Schmidt et al., 1986) to 10+ years (McDaniel, 1986). Reeve and Bonaccio (2011) reached a similar conclusion in their own

review of the relationship between cognitive ability and job performance, noting that “although validities might degrade somewhat over long intervals, we found no evidence to suggest that they degrade appreciably, thereby retaining practically useful levels of validity over very long intervals” (p. 269). Our analysis of the JPM data provide new support for this conclusion. Nevertheless, it remains possible that the validity of cognitive ability would drop to near zero over longer spans of time than have been examined in research (e.g., 20 years).

Discussion

What can be concluded about the role of cognitive ability in expertise? Table 2 summarizes findings from the expertise literature most directly relevant to the possibility of expertise-related mitigation of cognitive ability effects. These studies tested (or reported information to test) whether domain-specific factors mitigate effects of cognitive ability factors on domain-relevant performance, by either comparing ability-performance correlations across skill groups or testing interactions between domain-specific factors and cognitive ability factors. As shown, three studies provide evidence for expertise-related mitigation of cognitive ability effects and 10 studies do not; the results of two other studies are mixed or unclear. Based on sample size alone, the Burgoyne et al. (2016) chess meta-analysis might be seen as the best evidence for mitigation, but we

reiterate that expertise level (i.e., ranked or unranked) was highly confounded with both age and type of skill measure. This evidence certainly does not warrant any strong conclusions about expertise-related mitigation of the effects of cognitive ability.

Domain-General Models of Expertise: The Role of Cognitive Ability

Table 2 Summary of evidence for expertise-related mitigation of cognitive ability effects

Study	Domain	Sample Size		Cognitive Factor ^b	Evidence for Mitigation?	Empirical Test ^c
		<i>N</i>	Group <i>n</i> ^a			
Ceci & Liker (1986)	Handicapping	30	16/14	IQ	Unclear	Correlations
Masunaga and Horn (2001)	Go	263	62/89/92/23	Gf, Gc, PS	No	Correlations
Morrow et al. (2001)	Aviation	182	96/86	WMC, PS, VS	No	Interaction
Gallagher et al. (2003)	Surgery	98	48, 32/18	VS	No	Correlations
Sohn & Doane (2003)	Aviation	50	25/25	WMC	Yes	Interaction
Wanzel et al. (2003)	Surgery	47	27/12/8	VS	No	Correlations
Keehner et al. (2004)	Surgery	93	48/45	VS	No	Correlations

Domain-General Models of Expertise: The Role of Cognitive Ability

Sohn & Doane (2004)	Aviation	52	25/27	WMC	Mixed	Correlations
Taylor et al. (2005)	Aviation	97	25/53/19	WMC, PS, AC	No	Interaction
Enochsson et al. (2006)	Surgery	29	18/11	VS	No	Correlations
Ruthsatz et al. (2008)	Music	261	178/19/64	Gf	No	Correlations
Meinz & Hambrick (2010)	Music	57	NA	WMC	No	Interaction
Meinz et al. (2012)	Poker	155	NA	WMC	No	Interaction
Hambrick et al. (2012)	Geology	67	NA	VS	Yes	Interaction
Burgoyne et al. (2016) ^d	Chess	1,604	1,267/337	Gf	Yes	Correlations

Note. Studies are listed in chronological order.

Domain-General Models of Expertise: The Role of Cognitive Ability

(^a) The skill group *n* values are listed in order of increasing expertise. NA for skill group *n* indicates that expertise was treated only as a continuous variable.

(^b) Gf, fluid intelligence; Gc, crystallized intelligence; WMC, working memory capacity; PS, perceptual speed; VS, visuospatial ability; AC, attentional control.

(^c) In the *interaction* test, mitigation is tested by evaluating the statistical interaction between a domain-specific factor and a cognitive ability factor. In the *correlations* test, mitigation is tested by testing for a difference in correlations between a cognitive ability factor and performance across groups representing different levels of a domain-specific factor.

(^d) Meta-analysis.

What can be made of these results? Unfortunately, not much, because studies in the expertise literature differ in methodological/design characteristics such as sample size, type of criterion task, tests used to measure cognitive ability factors, and use of single versus composite measures to index cognitive ability factors. Any (or all) of these differences could explain differences across studies in cognitive ability–performance relationships. With some of our own studies as examples (e.g., *Meinz & Hambrick, 2010, N = 57*), it is worth emphasizing that sample sizes in this literature are often very small for research on individual differences (see Table 2), leading not only to low statistical power but low precision. Consequently, it is not surprising when results do not replicate (it is, in fact, often more surprising when they do). Note also that correlations in the expertise literature are seldom corrected for measurement error and/or restriction of range, resulting in systematic underestimates of the true magnitude of underlying relationships (see *McAbee & Oswald, 2017*).

A more consistent picture emerges from large-scale studies of job performance. Though validity may drop somewhat initially, measures of cognitive ability significantly predict job performance well beyond initial training. Expertise research often focuses on a specific aspect or component of performance in a domain (e.g., flight path prediction, poker hand evaluation); job performance research more often uses global measures of performance (e.g., overall supervisory ratings, total work sample scores). It could be that involvement of cognitive ability factors decreases as a function of skill in some components of a complex task or job but not in others (e.g., consistent but not variable components; *Ackerman, 1992*). This is one possible explanation for why correlations between cognitive ability and job performance may drop somewhat with job experience but still remain statistically significant.

Before proceeding, we note that when cognitive ability and domain-specific factors are measured in the same study, the latter generally account for more variance in expertise than the former (see *Ward et al., 2017*, for examples). At the same time, cognitive ability and domain-specific knowledge cannot generally be assumed to be independent. For example, *Schmidt, Hunter, and Outerbridge (1986)* found a correlation of 0.46 between AFQT scores and job knowledge. One interpretation of this finding is that measures of cognitive ability (e.g., IQ, working memory capacity) capture basic mental processes involved in acquiring information in learning situations (*Ackerman, 1996; Cattell, 1971; Jensen, 1998*). Moreover, even if domain-specific factors explain far more of the variance in expertise than domain-general factors, this does not preclude the latter from being practically useful. We examine this issue next.

Potential Uses of Cognitive Ability Measures to Accelerate Acquisition of Expertise

There are two major ways that cognitive ability measures might be used in efforts to accelerate acquisition of expertise. The first is for personnel selection and classification. That is, cognitive ability measures might be used to make hiring decisions and to assign employees to jobs once hired. The second area of application is in the design of training programs. If a certain cognitive ability factor (e.g., attentional control) is found to be a significant predictor of performance in a domain, then designing training to augment or *bootstrap* that ability (e.g., prompts to direct attention to task-relevant information) might be particularly beneficial for individuals lower in the ability (though see Hoffman et al., 2014, for a cautionary note about removing *desirable difficulties* from training).

But how large must a validity coefficient for a cognitive ability test be to justify its use for these applications? What qualifies as a practically significant effect? Given real-world outcomes (vs outcomes that do not generalize easily to the real world), moderate correlations can prove to be very important. Moreover, although variance explained (r^2) may be of theoretical interest to researchers (e.g., Macnamara, Hambrick, & Oswald, 2014), it is r and not r^2 that is an index of the direct relationship or the utility of a measure in terms of prediction (see Schmidt, Hunter, McKenzie, & Muldrow, 1979). As Kuncel and Hezlett (2010) commented:

Moderate relationships between predictors and criteria often are inappropriately discounted. For example, correlations of .30 have been dismissed as accounting for less than 10% of the variance in the criteria. However, this relationship is sufficiently large that hiring or admitting individuals who score better on the test can double the rate of successful performance. (p. 340)

This point was made nearly 80 years ago by Taylor and Russell (1939), who noted that interpreting the practical importance of correlation coefficients based on methods involving r^2

has led to some unwarranted pessimism on the part of many persons concerning the practical usefulness in an employment situation of validity coefficients in the range of those usually obtained. We believe that it may be of value to point out the very considerable improvement in selection efficiency which may be obtained with small correlation coefficients. (p. 571)

To that end, Taylor and Russell published a set of easy-to-use tables to determine the benefits of using selection tests of different validities in employment settings (see Law & Myers, 1993, for an automated approach). Three pieces of information are needed to use the tables: (1) the base rate of success in a job (i.e., the proportion of people who currently succeed in a job), (2) the selection ratio for the job (i.e., the ratio of applicants

who are selected), and (3) the validity of the test. With these three pieces of information, one can consult a Taylor–Russell table and find the predicted improvement in using the test for selection versus not using it.

Domain-General Models of Expertise: The Role of Cognitive Ability

Table 3 Example of Taylor-Russell Utility Table

Base Rate of Success = 0.20									
Selection Ratio ^a									
Validity (<i>r</i>)	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.00	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
0.10	0.25	0.24	0.23	0.23	0.22	0.22	0.21	0.21	0.21
0.20	0.31	0.28	0.27	0.26	0.25	0.24	0.23	0.22	0.21
0.30	0.37	0.33	0.30	0.28	0.27	0.25	0.24	0.23	0.21
0.40	0.44	0.38	0.34	0.31	0.29	0.27	0.25	0.23	0.22
0.50	0.52	0.44	0.38	0.35	0.31	0.29	0.26	0.24	0.22
0.60	0.60	0.50	0.43	0.38	0.34	0.30	0.27	0.24	0.22
0.70	0.69	0.56	0.48	0.41	0.36	0.31	0.28	0.25	0.22
0.80	0.79	0.64	0.53	0.45	0.38	0.33	0.28	0.25	0.22

Domain-General Models of Expertise: The Role of Cognitive Ability

0.90	0.91	0.75	0.60	0.48	0.40	0.33	0.29	0.25	0.22
-------------	------	------	------	------	------	------	------	------	------

Note. Validity (r): correlation between predictor variable and criterion variable.

(^a) Selection ratio: proportion of applicants who are hired. Values in the cells of the table indicate incremental validity, i.e., expected rate of success as a result of using the selection test.

Table 3 gives an example where the base rate is 0.20. As shown, when the selection ratio is low, even a selection test with modest validity will lead to a substantial improvement in employee performance. For example, if the selection ratio is 0.10, use of a test with validity of 0.20 to select applicants will lead to an 11 percent improvement over not using the test (or 17 percent for a test with validity of 0.30). However, if the selection ratio is high, even a test with high validity will yield little benefit. For example, if the selection ratio is 0.90, use of a test with validity of 0.80 will lead to an improvement of only 2 percent. More sophisticated approaches to utility analysis have been developed since Taylor and Russell published their tables (see Hunter & Schmidt, 1996; Schmidt, Hunter, Outerbridge, & Trattner, 1986), but suffice it to say that use of a test with a moderate level of validity can be practically useful.

Rosenthal and Rubin's (1982) binomial effect size display (BESD) provides another way to indicate the practical significance of an effect size of a given magnitude (see Ward et al., 2007, for an example of how the BESD can be used in expertise research). Displaying the difference between two proportions (e.g., treatment vs no treatment; selection test vs no selection test), like the Taylor-Russell tables the BESD reveals that modest effect sizes can be practically important. For example, Rosenthal (2005) explained that "an r of 0.20 is said to account for 'only 4% of the variance', but the BESD shows that this proportion of variance accounted for is equivalent to increasing the success rate ... from 40 to 60%." Figure 5 illustrates this point in terms of a hypothetical scenario where 100 individuals in an organization must be selected for a training program. In one case, a selection test with validity of 0.20 is used; in the other case, it is not used. As shown, using the selection test increases the chances that a trainee will pass the training program by 20 percent (i.e., 20 more people out of 100 pass), even though the scores on the test account for only 4 percent of the variance in the outcome.

Test Used to Select?	Training Outcome		Total
	Fail	Pass	
No	60	40	100
Yes	40	60	100
Total	100	100	

Figure 5 Example of binomial effect size display (BESD) relevant to expertise research and application.

Ethical Considerations

There are ethical issues associated with use of any psychological test to make decisions that affect people's lives (e.g., hiring decisions). Probably everyone would agree that it is unethical (not to mention legally unwise) to select individuals using a test with no demonstrated validity, but consider a situation where a test has modest validity—say, 0.30. One might argue that because the validity coefficient is far from perfect, it is unethical to use the test for selection because a considerable number of people with lower scores would be expected to succeed. However, one might also argue that *not* using the test for selection is unethical because lower-scoring individuals will be at a relatively high risk for failure, which may have adverse consequences for the individual (e.g., negative perceptions of other employees, lowered self-efficacy) and also the organization. Along with conducting a proper job analysis and validity study, any organization wishing to use a cognitive test for making personnel decisions must consider these sorts of ethical questions before putting the test into use (Landy & Conte, 2013).

Conclusions

Psychologists have long been interested in identifying traits that may help to explain individual differences in expertise (Hambrick, Campitelli, & Macamara, 2017). Here, we reviewed evidence for the contribution of cognitive ability. There is ample evidence that cognitive ability positively predicts individual differences in complex task performance early in training, but it is unclear whether it remains predictive after extensive practice or training. Evidence from research on traditional domains for expertise research (e.g., chess, music) is inconsistent. For some tasks, domain-specific factors may mitigate the effect of cognitive ability factors on performance (e.g., maintaining situational awareness in aviation), but for other tasks, this may not be the case (e.g., sight-reading music). Evidence from research on job performance is more consistent in indicating that measures of cognitive ability are predictive of job performance, well beyond initial training. In light of this evidence, we believe that at a broad level combining optimal procedures for training complex skills (Hoffman et al., 2014) with valid selection procedures holds tremendous promise for accelerating acquisition of expertise.

At a theoretical level, we believe that it is imperative for expertise researchers to develop and test formal models of expertise. Research on the involvement of cognitive ability factors in expertise has often proceeded somewhat haphazardly, with no general theory describing how mechanisms underlying performance differ across domains. There is no better illustration of this critical point than our own work. We have conducted a number of *one-off* studies—one in piano sight-reading (Meinz & Hambrick, 2010), another in geological bedrock mapping (Hambrick et al., 2012), another in Texas Hold 'em poker (Meinz et al., 2012)—with no theory to account for how results differ across these domains. Moving ahead, theories of expertise should draw on existing theoretical frameworks to identify potential predictors of expertise (e.g., Ackerman, 1996; Ericsson et al., 1993; Gagné, 2017). However, guided by both computational models (e.g., Altmann, Trafton, & Hambrick, 2014) and cognitive task analysis (Chipman, Schraagen, & Shalin, 2000), they must also specify the information processing mechanisms underlying performance in different types of tasks. Otherwise, there will continue to be no solid basis for comparing results across tasks, and evidence will remain fragmentary.

In the spirit of Hoffman et al.'s (2014) recommendations, we believe that it is also critical that expertise research expand beyond highly constrained activities such as chess, music, and sports, to *messy* real-world tasks in which the requirements of a job can change rapidly with technological developments and there is no well-circumscribed body of knowledge (as there is in, say, chess). We think that measures of cognitive ability factors hypothesized to underlie *adaptability* (e.g., attentional control, working memory capacity) may have particular promise for predicting performance in jobs such as these. These measures are also attractive because some research has suggested they may reduce group differences (e.g., by race/ethnicity) and resultant adverse impact in selection while still achieving high validity (Verive & McDaniel, 1996). More generally, we are optimistic

that the scientific knowledge that will accumulate through programmatic research on individual differences in expertise has great potential to inform efforts to accelerate the acquisition of societally important skills.

Acknowledgements

We thank Paul Ward and Jan Maarten Schraagen for their comments on an earlier version of this chapter.

References

- Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General* 117, 288–318.
- Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology* 77, 598–614.
- Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence* 22, 227–257.
- Altmann, E. M., Trafton, J. G., & Hambrick, D. Z. (2014). Momentary interruptions can derail the train of thought. *Journal of Experimental Psychology: General* 143, 215–226.
- Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science* 18, 89–94.
- Bayer, A. E., & Folger, J. (1966). Some correlates of a citation measure of productivity in science. *Sociology of Education* 39, 381–390.
- Berri, D. J., & Simmons, R. (2011). Catching a draft: On the process of selecting quarterbacks in the National Football League amateur draft. *Journal of Productivity Analysis* 35, 37–49.
- Burgoyne, A. P., Sala, G., Gobet, F., Macnamara, B. N., Campitelli, G., & Hambrick, D. Z. (2016). The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. *Intelligence* 59, 72–83.
- Burgoyne, A. P., Sala, G., Gobet, F., Macnamara, B. N., Campitelli, G., & Hambrick, D. Z. (in press). Corrigendum: The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. *Intelligence*.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Oxford, UK: Houghton Mifflin.

- Causse, M., Dehais, F., & Pastor, J. (2011). Executive functions and pilot characteristics predict flight simulator performance in general aviation pilots. *International Journal of Aviation Psychology* 21, 217-234.
- Ceci, S. J., & Liker, J. K. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General* 115, 255-266.
- Chipman, S. F., Schraagen, J. M., & Shalin, V. L. (2000). Introduction to cognitive task analysis. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds), *Cognitive task analysis* (pp. 3-23). New York: Lawrence Erlbaum.
- Clark, M. J., & Centra, J. A. (1982). *Conditions influencing the career accomplishments of Ph. Ds.* ETS Research Report Series, 1982.
- Creager, J. A., & Harmon, L. R. (1966). *On-the-job validation of selection variables*. Office of Scientific Personnel, National Academy of Sciences-National Research Council.
- Deary, I. J., Graham, K. S., & Maran, A. G. (1992). Relationships between surgical ability ratings and spatial abilities and personality. *Journal of the Royal College of Surgeons of Edinburgh* 37, 74-79.
- de Bruin, A. B., Kok, E. M., Leppink, J., & Camp, G. (2014). Practice, intelligence, and enjoyment in novice chess players: A prospective study at the earliest stage of a chess career. *Intelligence* 45, 18-25.
- de Groot, A.D. (1965/1978). *Thought and choice in chess*. Amsterdam: Amsterdam Academic Archive.
- Detterman, D. K. (2014). Introduction to the intelligence special issue on the development of expertise: Is ability necessary? *Intelligence* 45, 1-5.
- Detterman, D. K., & Spry, K. M. (1988). Is it smart to play the horses? Comment on "A day at the races: A study of IQ, expertise, and cognitive complexity" (Ceci & Liker, 1986). *Journal of Experimental Psychology: General* 117, 91-95.
- Engelhardt, L. E., Mann, F. D., Briley, D. A., Church, J. A., Harden, K. P., & Tucker-Drob, E. M. (2016). Strong genetic overlap between executive functions and intelligence. *Journal of Experimental Psychology: General* 145, 1141-1159.
- Enochsson, L., Westman, B., Ritter, E. M., Hedman, L., Kjellin, A., Wredmark, T., & Felländer-Tsai, L. (2006). Objective assessment of visuospatial and psychomotor ability and flow of residents and senior endoscopists in simulated gastroscopy. *Surgical Endoscopy and Other Interventional Techniques* 20, 895-899.
- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J.

Feltovich, & R. R. Hoffman (Eds), *The Cambridge handbook of expertise and expert performance* (pp. 683–703). New York: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review* 102, 211–245.

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

Ericsson, K. A., & Smith, J. (1991). *Toward a general theory of expertise: Prospects and limits*. Cambridge, UK: Cambridge University Press.

Farrell, J. N., & McDaniel, M. A. (2001). The stability of validity coefficients over time: Ackerman's (1988) model and the General Aptitude Test Battery. *Journal of Applied Psychology* 86, 60–79.

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

Fleishman, E. A., & Rich, S. (1963). Role of kinesthetic and spatial-visual abilities in perceptual-motor learning. *Journal of Experimental Psychology* 66, 6–11.

Folger, J. K., Astin, H. S., & Bayer, A. E. (1970). *Human resources and higher education: Staff report on the commission on human resources and advanced education*. Russell Sage Foundation.

Gagné, F. (2017). Expertise development from an IMTD perspective. In D. Z. Hambrick, G. Campitelli, & B. N. Macnamara (Eds), *The science of expertise: Behavioral, neural, and genetic approaches to complex skill* (pp. 307–327). New York: Routledge.

Gallagher, A. G., Cowie, R., Crothers, I., Jordan-Black, J. A., & Satava, R. M. (2003). PicSOR: an objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopic performance. *Surgical Endoscopy* 17, 1468–1471.

Gibbons, R. D., Baker, R. J., & Skinner, D. B. (1986). Field articulation testing: a predictor of technical skills in surgical residents. *Journal of Surgical Research* 41, 53–57.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence* 24, 13–23.

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

Groenier, M., Schraagen, J. M. C., Miedema, H. A., & Broeders, I. A. J. M. (2014). The role of cognitive abilities in laparoscopic simulator training. *Advances in Health Sciences Education* 19, 203–217.

Haier, R. J. (2016). *The neuroscience of intelligence*. Cambridge, UK: Cambridge University Press.

Hambrick, D. Z., & Campitelli, G., & Macamara, B. N. (2017). Introduction: A brief history of the science of expertise and overview of the book. In D. Z. Hambrick, G. Campitelli, & B. N. Macamara (Eds.), *The science of expertise: Behavioral, neural, and genetic approaches to complex skill* (pp. 1–10). New York: Routledge.

Hambrick, D. Z., Libarkin, J. C., Petcovic, H. L., Baker, K. M., Elkins, J., Callahan, C. N., ..., & LaDue, N. D. (2012). A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. *Journal of Experimental Psychology: General*, 141, 397–403.

Hambrick, D. Z., & Meinz, E. J. (2011). Limits on the predictive power of domain-specific experience and knowledge in skilled performance. *Current Directions in Psychological Science* 20, 275–279.

Hoffman, R. R., Ward, P., Feltovich, P. J., DiBello, L., Fiore, S. M., & Andrews, D. H. (2014). *Accelerated expertise: Training for high proficiency in a complex world*. New York: Psychology Press.

Hulin, C. L., Henry, R. A., & Noon, S. L. (1990). Adding a dimension: Time as a factor in the generalizability of predictive relationships. *Psychological Bulletin* 107, 328–340.

Hunter, J. E., & Schmidt, F. L. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law* 2, 447–472.

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

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

Kane, M. J., Conway, A. R., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, J. N. Towse (Eds), *Variation in Working Memory* (pp. 21–48). New York: Oxford University Press.

Kaufman, H. G. (1972). Relations of ability and interest to currency of professional knowledge among engineers. *Journal of Applied Psychology* 56, 495–499.

Keehner, M. M., Tendick, F., Meng, M. V., Anwar, H. P., Hegarty, M., Stoller, M. L., & Duh, Q. Y. (2004). Spatial ability, experience, and skill in laparoscopic surgery. *American Journal of Surgery* 188, 71–75.

Knopik, V. S., Neiderhiser, J. M., DeFries, J. C., & Plomin, R. (2016). *Behavioral genetics* (7th edn). New York: Macmillan Higher Education.

- Kopiez, R., & In Lee, J. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research*10, 41–62.
- Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science* 315(5815), 1080–1081.
- Kuncel, N. R., & Hezlett, S. A. (2010). Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current Directions in Psychological Science*19, 339–345.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance. *Psychological Bulletin*127, 162–181.
- Landy, F. J., & Conte, J. M. (2013). Work in the 21st century: An introduction to industrial and organizational psychology (4th edn). London: Wiley.
- Law, K. S., & Myors, B. (1993). Cutoff scores that maximize the total utility of a selection program: Comment on Martin and Raju's (1992) procedure. *Journal of Applied Psychology*78, 736–740.
- Lee, B., Park, J. Y., Jung, W. H., Kim, H. S., Oh, J. S., Choi, C. H., ..., & Kwon, J. S. (2010). White matter neuroplastic changes in long-term trained players of the game of "Baduk" (GO): a voxel-based diffusion-tensor imaging study. *Neuroimage*52, 9–19.
- Leonard, C. A., & Williams, R. J. (2015). Characteristics of good poker players. *Journal of Gambling Issues* 31, 45–68.
- Louridas, M., Szasz, P., de Montbrun, S., Harris, K. A., & Grantcharov, T. P. (2016). Can We predict technical aptitude?: A systematic review. *Annals of Surgery*263, 673–691.
- Lubinski, D. (2009). Exceptional cognitive ability: the phenotype. *Behavior Genetics* 39, 350–358.
- Lyons, B. D., Hoffman, B. J., & Michel, J. W. (2009). Not much more than g? An examination of the impact of intelligence on NFL performance. *Human Performance*22, 225–245.
- McAbee, S. T., & Oswald, F. L. (2017). Primer—Statistical methods in the study of expertise: Some key considerations. In D. Z. Hambrick, G. Campitelli, & B. N. Macnamara (Eds), *The science of expertise: Behavioral, neural, and genetic approaches to complex skill* (pp. 13–30). New York: Routledge.
- McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*24, 222–243.

McDaniel, M. A. (1986). The evaluation of a causal model of job performance: The interrelationships of general mental ability, job experience, and job performance (Unpublished doctoral dissertation). George Washington University, Washington, DC.

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.

Mann, D. T., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology*29, 457–478.

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

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.

Meinz, E. J., Hambrick, D. Z., Hawkins, C. B., Gillings, A. K., Meyer, B. E., & Schneider, J. L. (2012). Roles of domain knowledge and working memory capacity in components of skill in Texas Hold'Em poker. *Journal of Applied Research in Memory and Cognition*1, 34–40.

Morrow, D. G., Menard, W. E., Ridolfo, H. E., Stine-Morrow, E. A. L., Teller, T., & Bryant, D. (2003). Expertise, cognitive ability, and age effects on pilot communication. *International Journal of Aviation Psychology* 13, 345–371.

Morrow, D. G., Menard, W. E., Stine-Morrow, E. A., Teller, T., & Bryant, D. (2001). The influence of expertise and task factors on age differences in pilot communication. *Psychology and Aging*16, 31–46.

Murdoch, J. R., Bainbridge, L. C., Fisher, S. G., & Webster, M. H. (1994). Can a simple test of visual-motor skill predict the performance of microsurgeons? *Journal of the Royal College of Surgeons of Edinburgh*39, 150–152.

Petushek, E. J. (2014). Development and validation of the anterior cruciate ligament injury-risk-estimation quiz (ACL-IQ). Retrieved from ProQuest Dissertations & Theses Global. (Accession No. 3643834).

Petushek, E. J., Cokely, E. T., Ward, P., & Myer, G. D. (2015). Injury risk estimation expertise: cognitive-perceptual mechanisms of ACL-IQ. *Journal of Sport and Exercise Psychology*37, 291–304.

Pfau, H. D., & Murphy, M. D. (1988). Role of verbal knowledge in chess skill. *American Journal of Psychology*101, 73–86.

- Reeve, C. L., & Bonaccio, S. (2011). On the myth and the reality of the temporal validity degradation of general mental ability test scores. *Intelligence*39, 255–272.
- Risucci, D., Geiss, A., Gellman, L., Pinard, B., & Rosser, J. (2001). Surgeon-specific factors in the acquisition of laparoscopic surgical skills. *American Journal of Surgery*181, 289–293.
- Rosenthal, R. (2005). Binomial effect size display. *Encyclopedia of statistics in behavioral science*. Wiley StatsRef: Statistics Reference Online.
- Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*74, 166–169.
- 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.
- Sala, G., Burgoyne, A. P., Macnamara, B. N., Hambrick, D. Z., Campitelli, G., & Gobet, F. (2017). Checking the “Academic Selection” argument. Chess players outperform non-chess players in cognitive skills related to intelligence: A meta-analysis. *Intelligence*61, 130–139.
- Schellenberg, E. G., & Weiss, M. W. (2013). Music and cognitive abilities. In D. Deutsch (Ed.), *Psychology of music* (3rd edn, pp. 499–550). London: Academic Press.
- Schmidt, F. L., Hunter, J. E., McKenzie, R. W., & Muldrow, T. W. (1979). Impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology* 64, 609–626.
- 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.
- 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.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology* 71, 432–439.
- Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Goff, S. (1988). Joint relation of experience and ability with job performance: Test of three hypotheses. *Journal of Applied Psychology*73, 46–57.
- Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Trattner, M. H. (1986). The economic impact of job selection methods on size, productivity, and payroll costs of the federal work force: An empirically based demonstration. *Personnel Psychology* 39, 1–29.

- Schueneman, A. L., Pickleman, J., Hesslein, R., & Freeark, R. J. (1984). Neuropsychologic predictors of operative skill among general surgery residents. *Surgery*96, 288-295.
- Sohn, Y. W., & Doane, S. M. (2003). Roles of working memory capacity and long-term working memory skill in complex task performance. *Memory & Cognition* 31, 458-466.
- Sohn, Y. W., & Doane, S. M. (2004). Memory processes of flight situation awareness: Interactive roles of working memory capacity, long-term working memory, and expertise. *Human Factors*46, 461-475.
- Takagi, H. (1997). *Cognitive aging: expertise and fluid intelligence* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 9733148).
- Taylor, J. L., O'Hara, R., Mumenthaler, M. S., Rosen, A. C., & Yesavage, J. A. (2005). Cognitive ability, expertise, and age differences in following air-traffic control instructions. *Psychology and Aging*20, 117-133.
- Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. *Journal of Applied Psychology*23, 565-578.
- Toma, M., Halpern, D. F., & Berger, D. E. (2014). Cognitive abilities of elite nationally ranked SCRABBLE and crossword experts. *Applied Cognitive Psychology*28, 727-737.
- Tucker-Drob, E. M. (2011). Global and domain-specific changes in cognition throughout adulthood. *Developmental Psychology*47, 331-343.
- 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.
- Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*9, 160-164.
- Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2015). Working memory delay activity predicts individual differences in cognitive abilities. *Journal of Cognitive Neuroscience* 27, 853-865.
- Unterrainer, J. M., Kaller, C. P., Halsband, U., & Rahm, B. (2006). Planning abilities and chess: A comparison of chess and non-chess players on the Tower of London task. *British Journal of Psychology*97, 299-311.
- Unterrainer, J. M., Kaller, C. P., Leonhart, R., & Rahm, B. (2011). Revising superior planning performance in chess players: the impact of time restriction and motivation aspects. *American Journal of Psychology*124, 213-225.
- Verive, J. M., & McDaniel, M. A. (1996). Short-term memory tests in personnel selection: Low adverse impact and high validity. *Intelligence* 23, 15-32.

Vestberg, T., Gustafson, R., Maurex, L., Ingvar, M., & Petrovic, P. (2012). Executive functions predict the success of top-soccer players. *PloS one*7, e34731.

Voss, M. W., Kramer, A. F., Basak, C., Prakash, R. S., & Roberts, B. (2010). Are expert athletes “expert” in the cognitive laboratory? A meta-analytic review of cognition and sport expertise. *Applied Cognitive Psychology*24, 812–826.

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

Ward, P., Belling, P., Petushek, E., & Ehrlinger, J. (2017). Does talent exist? A re-evaluation of the nature–nurture debate. In J. Baker, S. Cobley, J. Schorer, & N. Wattie (Eds), *Routledge handbook of talent identification and development in sport* (pp. 19–34). New York: Routledge.

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.

Wigdor, A. K., & Green, B. F. (1991). *Performance assessment for the workplace*, Volumes I & II. Washington, DC: National Academy Press.

Notes:

(1.) We thank Michael McDaniel for sending us a copy of this study.

(2.) We are grateful to Dr. Jane M. Arabian (Assistant Director, Accession Policy Office of the Under Secretary of Defense, The Pentagon, Washington, DC) for granting us permission to use the data, and to Dr. Rodney McCloy (Principal Scientist, Human Resources Research Organization, Louisville, KY) for sending us the data, with helpful notes.

(3.) Prior to conducting statistical analyses, we screened the variables for values greater than $|3.5|$ SDs from the total sample mean (i.e., univariate outliers); 31 of the 30,264 values (0.1%) met this criterion, and we truncated these values to the $|3.5|$ SD cutoff value. There was one participant with zero months of job experience; prior to log transforming the job experience variable, we set this value to 0.03 months (1 day).

David Z. Hambrick

David Z. Hambrick, Department of Psychology, Michigan State University, USA

Alexander P. Burgoyne

Alexander P. Burgoyne, Department of Psychology, Michigan State University, USA

Domain-General Models of Expertise: The Role of Cognitive Ability

Frederick L. Oswald

Frederick L. Oswald, Department of Psychology, Rice University, USA

