Journal of Applied Psychology

ISSN: 0021-9010

2024, Vol. 109, No. 3, 437–455
https://doi.org/10.1037/apl0001150

RESEARCH REPORT

The Validity of General Cognitive Ability Predicting Job-Specific Performance Is Stable Across Different Levels of Job Experience

David Z. Hambrick[1], Alexander P. Burgoyne[2], and Frederick L. Oswald[3]
[1] Department of Psychology, Michigan State University
[2] School of Psychology, Georgia Institute of Technology
[3] Department of Psychological Sciences, Rice University

Decades of research in industrial–organizational psychology have established that measures of general cognitive ability (*g*) consistently and positively predict job-specific performance to a statistically and practically significant degree across jobs. But is the validity of *g* stable across different levels of job experience? The present study addresses this question using historical large-scale data across 31 diverse military occupations from the Joint-Service Job Performance Measurement/Enlistment Standards Project ($N = 10{,}088$). Across all jobs, results of our meta-analysis find near-zero interactions between Armed Forces Qualification Test score (a composite of math and verbal scores) and time in service when predicting job-specific performance. This finding supports the validity of *g* for predicting job-specific performance even with increasing job experience and provides no evidence for diminishing validity of *g*. We discuss the theoretical and practical implications of these findings, along with directions for personnel selection research and practice.

Keywords: job performance, cognitive ability, job experience

Supplemental materials: https://doi.org/10.1037/apl0001150.supp

Research has established that measures of general cognitive ability (or *g*) correlate positively with reliable measures of job-specific performance (Schmidt & Hunter, 2004).[1] But does the validity of *g* remain stable with increasing job experience? Research addressing this question is relevant to a view that has been frequently expressed in both the scientific literature and popular press: *g* becomes less important as a predictor of performance as knowledge and skills accumulate through training. For example, in his book *Peak: Secrets from the New Science of Expertise*, the late cognitive psychologist K. Anders Ericsson claimed that "the influence of general abilities, such as intelligence quotient, is greater on performance of beginners but virtually disappears for individual differences among expert performers" (Ericsson & Pool, 2016, p. 708, see also Ericsson, 2018).

Schmidt et al. (1988) referred to this possibility as the *convergence hypothesis*, which predicts that the difference in job-specific performance between lower ability and higher ability groups decreases as job experience increases (Figure 1, left panel). They contrasted this hypothesis with two alternative hypotheses. The *divergence hypothesis* predicts that the performance differences between ability groups increases as job experience increases (Figure 1, center panel), whereas the *noninteractive hypothesis* predicts that performance differences remain constant with experience (i.e., main effects and No $g \times$ Job Experience interaction; Figure 1, right panel).

The convergence hypothesis has received some support from studies examining the effects of short periods of training (typically <5 hr) on performance in laboratory tasks. For example, Ackerman (1988) found that the correlation between a measure of *g* and performance on visual search tasks decreased when the stimulus–response mappings were always the same (see, e.g., Ackerman's Experiment 2, $N = 191$; Session 1 $r = .41$, 95% CI [.25, .77], to Session 12 $r \approx .08$, 95% CI [−.06, .22]; see also Fleishman & Rich, 1963). However, an obvious limitation of this type of study is that the period of training is much shorter than it is for most jobs. Keil and Cortina (2001) performed a meta-analysis that included studies examining the relationship between measures of *g* (e.g., SAT, Raven's matrices) and criterion measures reflecting longer periods of learning. Criterion tasks were coded as having requirements that were either inconsistent or consistent (Ackerman, 1988). Results pertaining to studies with a "long" time span (>1 day) are most relevant here. For inconsistent tasks, the correlation between *g* and performance was $r \approx .20$ at around 300 hr of learning, compared to $r \approx .10$ at around 700 hr. For consistent tasks, correlations decreased from $r \approx .40$, then leveled off at approximately 1 year, and then decreased again at 3.8 years, although the average correlation was still sizeable, $r \approx .20$, at the longest delay (5.7 years). It should be noted, however, that the results of this meta-analysis are somewhat difficult to interpret, given the

Editor's Note. In-Sue Oh served as the action editor for this article.—LTE

This article was published Online First October 16, 2023.

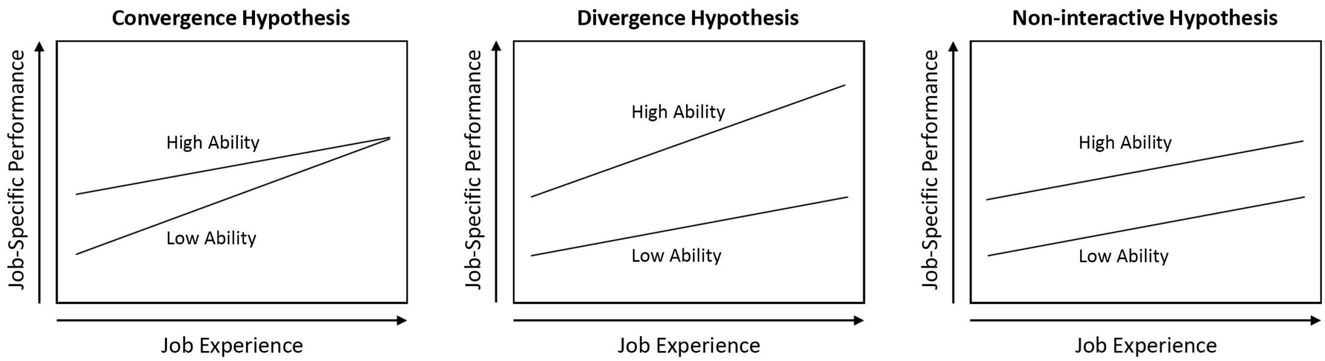
David Z. Hambrick https://orcid.org/0000-0002-5315-5597

Correspondence concerning this article should be addressed to David Z. Hambrick, Department of Psychology, Michigan State University, East Lansing, MI 48824, United States. Email: hambric3@msu.edu

[1] Job-specific performance is also known as *core task performance* or *core technical performance* (e.g., Campbell & Wiernik, 2015; Viswesvaran & Ones, 2002).

Figure 1

Alternative Hypotheses Concerning Joint Effects of General Cognitive Ability (g) and Job Experience on Job-Specific Performance



Note. The convergence hypothesis predicts an Underadditive Ability × Job Experience interaction. The divergence hypothesis predicts an Overadditive Ability × Job Experience interaction. The noninteractive hypothesis predicts main effects of ability and job experience, with no interaction. Adapted from "Joint Relation of Experience and Ability With Job Performance: Test of Three Hypotheses," by F. L. Schmidt, J. E. Hunter, A. N. Outerbridge, and S. Goff, 1988, *Journal of Applied Psychology*, *79*(1), p. 47 (https://doi.org/10.1037/0021-9010.73.1.46). Copyright 1988 by the American Psychological Association.

inclusion of not only measures of performance in specific jobs (e.g., butter wrapping; Rothe, 1946) but also broader measures, such as grade-point average reflecting academic performance.

Large-scale studies of job-specific performance provide more interpretable evidence concerning the joint effects of cognitive ability and job experience and most consistently have supported the noninteractive (main effects) hypothesis (see Hambrick et al., 2019). Using a U.S. Employment Service database consisting of a large and diverse sample of civilian workers, McDaniel (1986) found that correlations between a composite measure of g from the General Aptitude Test Battery and supervisory ratings of performance remained nearly constant as a function of job experience ($r = .23$, 95% CI [.17, .29] at <1 year, $n = 1{,}019$, to $r = .20$, 95% CI [.14, .26] at >10+ years, $n = 879$), consistent with the noninteractive hypothesis. Extending this work, Farrell and McDaniel (2001) classified the jobs by whether their task demands were consistent (fixed) or inconsistent (changing), using two different definitions of consistency: *complexity* (where more complex = less consistent) and *tolerance for repetition* (where more tolerance = more consistent). Under both definitions, g was a statistically and practically significant predictor of job-specific performance across levels of task consistency (avg. $r = .23$). Moreover, the level of validity was meaningful even at the highest levels of job experience (at 109–120 months, $r = .26$, 95% CI [.11, .40], $n = 207$, to $r = .34$, 95% CI [.21, .46], $n = 161$; avg. $r = .29$).

Additional support for the noninteractive hypothesis comes from large-scale studies of military personnel. A study by Schmidt et al. (1988) focused on U.S. Army soldiers ($N = 1{,}474$) in four jobs (armor repairman, armor crewman, supply specialist, and cook). Job-specific performance was measured using work samples and supervisor ratings. The measure of g was the Armed Forces Qualification Test (AFQT) score, which is a composite of math and verbal subtest scores from the Armed Services Vocational Aptitude Battery (ASVAB); job experience was operationalized as amount of experience in the specific job. Both AFQT and job experience had positive effects on job-specific performance. However, only one of 12 AFQT × Job Experience interactions (work sample performance for armor crewman) was statistically significant, but this could be attributed to chance. Further, this one interaction was not clearly interpretable as supporting the convergence of ability groups.

The present study makes use of data from another large-scale military personnel project: the Joint-Service Job Performance Measurement/Enlistment (JPM) Standards Project. The JPM project was conducted in the 1980s as part of an effort to improve military personnel retention and success and included data on military personnel across four U.S. military branches in a diverse set of 31 military occupation specialties (MOSs), such as medical specialist (Army), helicopter mechanic (Marines), air-traffic control operator (Air Force), and electronics technician (Navy). As with the Schmidt et al. (1988) study, the measure of g was the AFQT composite score, which comprises four verbal and math subtests of the ASVAB, which all tend to be highly correlated (generally $r = .60$ or higher). The measure of job experience was time in service (in months); job-specific performance was measured behaviorally via hands-on job performance tests (HOPTs). In these tests, a subject matter expert (e.g., a noncommissioned officer) rated the test-taker's performance on completing each of a series of job-relevant tasks (e.g., putting on a field dressing for a medic).

Collapsing across the 31 MOSs in the JPM data set, Wigdor and Green (1991) presented the mean level of job-specific performance across years in service for AFQT score quartiles. There was a positive association between g and job-specific performance, such that the difference in mean job-specific performance between the lowest ability group and the highest ability group at 0–1 year of service was approximately $d = 1.0$, but this difference remained quite large after 4+ years of service, at about $d = .50$. In another analysis of JPM data, focusing specifically on the jet mechanic MOS ($N = 255$), Lance et al. (1989) found that both a measure of mechanical aptitude from the ASVAB and time in service positively predicted scores on job-specific performance tests, although effect sizes were small (partial r^2s $< .015$). The Mechanical Aptitude × Time in Service interaction effect on job-specific performance was near zero (partial $r^2 < .001$). Focusing on the MOSs of automotive mechanic ($N = 891$) and helicopter mechanic ($N = 522$) in the JPM data, Mayberry and Carey (1997) presented evidence for an interaction between mechanical aptitude and a self-rating of task-specific experience predicting job-specific performance supporting the convergence hypothesis. However, there was no interaction between mechanical aptitude and time in service.

The Present Study

Taken together, these results support the noninteractive hypothesis of the joint effects of *g* and job experience on job-specific performance. Effect sizes for *g* and job experience are generally positive and moderate in magnitude. By contrast, when tested, the $g \times$ Job Experience interaction effect is typically found to be very small, and correlations between *g* and job-specific performance are most often moderate in magnitude even at the highest levels of job experience considered, supportive of Schmidt et al.'s (1988) noninteractive hypothesis. However, most studies examining this question have focused on lab tasks rather than actual jobs (e.g., Ackerman, 1988), or a single job or a small set of jobs in the field, limiting the generalizability of the findings. In Schmidt et al.'s (1988) study, only four MOSs were considered, and the Lance et al. (1989) and Mayberry and Carey (1997) studies focused only on mechanic MOSs.

Other relevant studies (e.g., Farrell & McDaniel, 2001) have used global supervisory ratings as the measure of job performance rather than more objective job-specific measures. Our research question extends this work by empirically examining the possibility that knowledge and skills acquired through job experience, across a diverse range of specific jobs, diminish the importance of *g* as a predictor of job-specific performance, which would support the convergence hypothesis. Whether in a research or applied context, global supervisory ratings of overall level of job performance may be useful for administrative purposes, such as when making a promotion decision. However, from the standpoint of our research question, global supervisory ratings are problematic because they may reflect not only an individual's knowledge and skill in some job task (i.e., job-specific performance) but also nonskill factors, such as organizational citizenship behaviors (Johnson, 2001). Thus, the JPM study is ideal for answering our research question because hands-on job performance protocols were specifically designed to measure individuals' knowledge and skill in specific job tasks. Wigdor and Green (1991) presented mean HOPT scores across years in service for AFQT quartiles in the JPM study. However, they did not directly evaluate Schmidt et al.'s (1988) noninteractive and interactive hypotheses, as we will, by testing for main and interactive effects of the predictor variables on hands-on job performance.

The present study goes beyond previous research by addressing all the preceding limitations: the study included a large sample of participants and a diverse range of actual jobs; the measures of job-specific performance captured objectively scored performance in specific job tasks; and we directly tested for main and interactive effects of AFQT and job experience on job-specific performance. Another strength of this study is that we integrated the results of our analyses using meta-analysis. Capitalizing on the large sample size of the JPM study, a meta-analytic approach allowed us to both aggregate MOS-specific findings and evaluate their heterogeneity. The results provide what are arguably the most informative empirical estimates to date of the interactive effect of *g* and job experience on job-specific performance, above and beyond the main effects of these predictor variables.

Research Questions

We addressed two research questions in this study. The first concerned the joint effects of *g* (as indexed by the AFQT math and verbal composite) and job experience (as indexed by time in service) on job-specific performance (as measured by HOPTs). Based on past research (e.g., Quiñones et al., 1995; Schmidt et al., 1988; Schmidt & Hunter, 2004), we predicted that the average correlation for both AFQT and time in service with job-specific performance would be positive, and both statistically and practically significant. The more interesting question was whether we would find a significant AFQT $\times$ Time in Service interaction, on average, and for each MOS. Using Figure 1 as a reference, a negative average interaction effect would support the convergence hypothesis; a positive average interaction effect would support the divergence hypothesis; and an average interaction effect near zero would support the noninteractive hypothesis (i.e., main effects only).

The second research question was whether the AFQT $\times$ Time in Service interaction would vary as a function of the type of job (MOS). In moderator analyses, we used job characteristic component scores from an analysis of the JPM data by McCloy (1994; see online Supplemental Materials). The four characteristics, reflecting the job requirements of the MOSs, were: (a) *working with things:* work requiring some form of manual labor; (b) *cognitive complexity:* work involving abstract activities (e.g., synthesizing data, mentoring people) versus more concrete activities (e.g., copying data, taking instructions, handling equipment); (c) *unpleasant working conditions:* work that is hazardous, fatiguing, or uncomfortable; and (d) *fine motor control*: work that requires the coordination of small muscle movements (e.g., knot tying, magazine loading). Here, we tested whether these four job characteristics moderated the effects of AFQT, time in service, and the AFQT $\times$ Time in Service interaction on job-specific performance.

We were primarily interested in potential moderating effects of cognitive complexity, which reflects task demands such as integrating information from multiple sources, solving novel problems, and holding information in working memory. The classic meta-analysis by Hunter and Hunter (1984) found that the validity of *g* for job-specific performance (correcting for criterion unreliability and range restriction) was larger for jobs higher in complexity (e.g., manager; $\rho = .53$) than for jobs lower in complexity (e.g., vehicle operator; $\rho = .28$). In the present study, we asked whether the level of cognitive complexity moderates the *interaction* of *g* and job experience on job-specific performance. Jobs lower in cognitive complexity (e.g., infantryman) should place lower demands on *g* than jobs higher in cognitive complexity (e.g., air-traffic controller). Moreover, higher complexity jobs should have a greater requirement for higher level cognitive processes, such as problem solving, reasoning, and comprehension, than lower complexity jobs. Accordingly, in lower complexity jobs, *g* might attenuate or drop out entirely as a predictor of performance differences after employees receive training. By contrast, higher complexity jobs might be expected to involve *g* even after extensive training.

It follows that cognitive complexity should moderate the AFQT $\times$ Time in Service interaction, such that the interaction effect will be negative for low complexity jobs (Figure 1, left panel), indicating convergence of ability groups with increasing training, but near zero for high complexity jobs (Figure 1, right panel), indicating preservation of the difference between ability groups. Alternatively, the interaction effect for high complexity jobs may even be positive (Figure 1, center panel), indicating that the difference between ability groups increases with training because higher ability individuals derive greater benefit from training.

We were also interested in possible effects of the working with things characteristic. Working with things reflects the extent to which jobs require manual (i.e., physical) labor. One possibility is that effects of g on job performance tend to be smaller for jobs high on working with things (e.g., infantry vs. air-traffic controller). If jobs that primarily involve manual labor can be performed more automatically, then increasing job experience might attenuate the effect of g on job performance in manual labor jobs to a greater degree than in other types of jobs. This speculation would be supported by finding an Underadditive $g \times$ Time in Service interaction. We had no specific hypotheses concerning the other job characteristics (unpleasant working conditions and fine motor control), so we tested for effects of the component scores reflecting those job characteristics on a purely exploratory basis.

To address the preceding research questions, for each of the 31 MOSs in the JPM study, we regressed the measure of job-specific performance onto AFQT (as an estimate of g), time in service (as an estimate of job experience), and the AFQT $\times$ Time in Service interaction term. Given that the metric was the same for both predictor and criterion variables across MOSs, we then were able to use meta-analysis to synthesize the regression results by (a) estimating these average effects across MOSs and (b) analyzing whether these regression effects were affected (moderated) by the aforementioned job characteristics.

Method

The Data Set

The JPM data set includes 10,088 military personnel from across the four U.S. military branches across 31 MOSs.[2] As shown in Table 1, the MOS sample sizes (number of personnel) ranged from 80 (Navy MOS of electrician's mate) to 940 (Marines MOS of rifleman), with a median N of 216. Approximately 92% of the total sample ($N = 9{,}261$) were high school graduates; the remaining 8% ($N = 827$) were nonhigh school graduates. Demographic statistics are available for subsamples of the total sample (see Wigdor & Green, 1991, Chapter 8) and indicate that the sample was predominantly male and White. In one subsample ($N = 7{,}044$), 21.1% of the personnel were classified as Black and 78.9% as nonminority; in another subsample ($N = 4{,}268$), 81% of the personnel were male and 19% were female.

Transparency and Openness

We describe all measures in the data set we obtained for the study and adhered to the *Journal of Applied Psychology* methodological checklist. Data are not available because we only obtained permission to use the data in the reported analyses (see Footnote 2, for more information on the source of the data). Data were analyzed using R, Version 3.63 (R Core Team, 2021), and the `metafor` meta-analysis package (Viechtbauer, 2010). Additional information on our methods, analysis code, and results can be found in the online Supplemental Materials and on the Open Science Framework (OSF) at https://osf.io/ap3mz/?view_only=641a709bb3394873aa70851db4b1edbb.

The study design, hypotheses, and analyses were not preregistered; the study was a secondary data analysis.

Measures

The AFQT score was used as a measure of g. The AFQT score is based on four subtests from the ASVAB: arithmetic reasoning (AR), mathematics knowledge (MK), paragraph comprehension (PC), and word knowledge (WK). The score is calculated as AFQT = 2 * (PC + WK) + AR + MK. Subtest scores were not available to us, and thus we could not compute a reliability estimate for the AFQT composite. However, in other work, the AFQT composite has been found to have good or better internal consistency and test–retest reliability (e.g., >.80 for both in Gade & Dudley, 2004; Palmer et al., 1988). The JPM data set made available to us included AFQT standard scores ($M = 207.52$, $SD = 21.74$); as described in more detail below, we converted these scores to percentile scores ($M = 54.31$, $SD = 19.99$) based on a nationally representative sample (the 1980 Profile of American Youth sample; Maier, 1985, p. v) so that we could correct correlations of AFQT with job-specific performance for direct range restriction in AFQT scores (see online Supplemental Materials, for further details).

Job experience was operationalized as time in service, in months. The measure of job-specific performance for each MOS was the HOPT score based on the percentage of MOS-specific task steps that the soldier was observed to perform successfully (for further information on the HOPT measures, see Wigdor & Green, 1991; Wise, 1994). These percentages can be compared across MOS, though there are important differences in the substance of the tasks performed and the number of tasks that were scored. Thus, differences in correlation and regression coefficients are partially attributable to differences in the performance criterion itself, which is why we examined component scores reflecting job characteristics as moderators of the ability–experience interaction.

Job tasks to be performed were selected on the basis that they were important to the MOS, frequently performed, not too dangerous, and challenging enough to impart variability in the total performance score and avoid floor/ceiling effects (Wise, 1994). For scoring purposes, the tasks were divided into multiple steps, and performance on each step was scored dichotomously as either correct or incorrect, with few exceptions. Although some steps may be seen as more critical than others, there were weighted equally in the determination of overall task performance. Referring to the measure of hands-on job performance, Wise (1994) stated: "It is not unreasonable to interpret these scores in a general sense as the percentage of the central or important parts of the job that the individual can perform successfully" (p. 42).

Component scores from McCloy's (1994) principal component analysis served as the measures of job characteristics (i.e., working with things, cognitive complexity, unpleasant working conditions, fine motor control; see online Supplemental Materials, for further information on this analysis). Table 1 reports each MOS's ratings on the job characteristics; Table A1 reports the correlations between mean MOS AFQT scores and time in service with the job characteristic ratings.

[2] We thank 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. We also thank Dr. Rodney McCloy (principal scientist, Human Resources Research Organization, Louisville, Kentucky) for sending us the data with helpful notes. Researchers interested in using the data set should contact Dr. Arabian at jane.m.arabian.civ@mail.mil.

Table 1
Descriptive Statistics for 31 MOSs

MOS	Description	*N* (%)	Job-specific performance *M* (*SD*)	AFQT *M* (*SD*)	Time in service *M* (*SD*)	Time in service Log *M* (*SD*)	Working with things	Cognitive complexity	Unpleasant working conditions	Fine motor control
Army		*4,941 (49)*	*71 (7)*	*52 (7)*	*21 (1)*	*1.30 (.02)*	*−0.64*	*−0.99*	*0.86*	*0.17*
11B	Infantryman	663 (6.6)	70 (7)	53 (22)	20 (5)	1.29 (.11)	−0.68	−2.41	2.33	0.18
13B	Cannon crewman	597 (5.9)	62 (11)	45 (19)	21 (6)	1.31 (.12)	−0.37	−2.61	1.72	−0.06
19E	Tank crewman	465 (4.6)	77 (8)	51 (21)	19 (5)	1.27 (.10)	−0.11	−1.95	1.50	0.15
31C	Single-channel radio operator	346 (3.4)	70 (8)	56 (20)	20 (5)	1.29 (.11)	0.17	0.18	−0.65	−0.16
63B	Light wheel vehicle mechanic	594 (5.9)	84 (5)	48 (19)	20 (5)	1.29 (.11)	0.49	0.32	0.94	0.16
64C	Motor transport operator	646 (6.4)	71 (8)	42 (19)	20 (5)	1.30 (.10)	−0.74	−2.00	0.49	0.49
71L	Administrative specialist	490 (4.9)	59 (8)	53 (19)	22 (5)	1.34 (.09)	−1.27	−0.60	−0.76	−0.42
91A	Medical specialist	483 (4.8)	71 (7)	59 (17)	21 (6)	1.31 (.11)	−1.60	0.12	1.02	1.36
95B	Military police	657 (6.5)	70 (6)	63 (15)	21 (5)	1.32 (.11)	−1.64	0.04	1.13	−0.14
Marines		*3,196 (32)*	*68 (12)*	*57 (6)*	*34 (8)*	*1.45 (.12)*	*0.23*	*−0.73*	*1.55*	*0.08*
031	Rifleman	940 (9.3)	53 (9)	49 (18)	23 (12)	1.29 (.25)	−0.68	−2.41	2.33	0.18
033	Machinegunner	271 (2.7)	55 (8)	50 (19)	30 (19)	1.40 (.27)	−0.55	−2.45	2.17	0.26
034	Mortarman	253 (2.5)	53 (9)	52 (19)	24 (17)	1.28 (.30)	−0.55	−2.45	2.17	0.26
035	Assaultman	277 (2.7)	64 (7)	58 (18)	29 (16)	1.38 (.27)	−0.82	−1.65	2.01	0.10
3521	Organizational automotive mechanic	907 (9.0)	77 (8)	54 (16)	39 (22)	1.52 (.25)	0.49	0.32	0.94	0.16
6112	Helicopter mechanic CH-46	152 (1.5)	79 (7)	63 (16)	41 (25)	1.54 (.27)	1.05	0.51	1.09	−0.05
6113	Helicopter mechanic CH-53	93 (0.9)	80 (9)	65 (17)	35 (23)	1.48 (.24)	1.05	0.51	1.09	−0.05
6114	Helicopter mechanic U/AH-1	190 (1.9)	72 (8)	64 (17)	38 (23)	1.52 (.24)	1.05	0.51	1.09	−0.05
6115	Helicopter mechanic CH-53 E	113 (1.1)	77 (7)	63 (17)	49 (24)	1.64 (.19)	1.05	0.51	1.09	−0.05
Air Force		*1,244 (12)*	*70 (7)*	*65 (10)*	*28 (3)*	*1.41 (.05)*	*0.02*	*−0.17*	*−0.22*	*0.18*
112	Aircrew life support specialist	166 (1.6)	70 (14)	59 (17)	29 (11)	1.42 (.20)	−0.11	−1.20	−0.44	0.08
272	Air traffic control operator	171 (1.7)	70 (10)	73 (15)	27 (9)	1.41 (.15)	−1.49	0.59	−0.59	0.56
324	Precision measuring equipment specialist	124 (1.2)	76 (9)	80 (13)	28 (11)	1.41 (.17)	0.43	0.01	0.28	0.23
328	Avionic communications specialist	83 (0.8)	77 (11)	81 (12)	35 (15)	1.50 (.21)	1.17	0.18	−0.35	−0.19
423	Aerospace ground equipment specialist	216 (2.1)	57 (10)	59 (17)	28 (10)	1.42 (.17)	0.09	−0.54	−0.12	0.93
426	Jet engine mechanic	188 (1.9)	73 (10)	56 (19)	29 (11)	1.44 (.17)	1.05	0.51	1.09	−0.05
492	Information systems radio operator	120 (1.2)	71 (13)	58 (19)	24 (13)	1.30 (.25)	0.17	0.18	−0.65	−0.16
732	Personnel specialist	176 (1.7)	62 (14)	58 (17)	28 (12)	1.40 (.21)	−1.18	−1.12	−0.95	0.06
Navy		*707 (7)*	*79 (6)*	*57 (15)*	*35 (5)*	*1.51 (.07)*	*0.66*	*0.14*	*−0.15*	*0.09*
EM	Electrician’s mate	80 (0.8)	81 (13)	54 (20)	34 (11)	1.50 (.16)	1.17	0.13	−0.17	−0.42
ET	Electronics technician	136 (1.3)	82 (9)	80 (14)	44 (13)	1.62 (.13)	0.56	0.39	−1.14	1.13
GSM	Gasoline turbine mechanic	89 (0.9)	81 (13)	48 (20)	33 (13)	1.49 (.15)	0.86	−0.20	0.42	−0.32
MM	Machinist’s mate	178 (1.8)	82 (9)	64 (23)	35 (14)	1.50 (.19)	0.86	−0.20	0.42	−0.32
RM	Radioman	224 (2.2)	67 (10)	42 (15)	31 (13)	1.45 (.19)	−0.17	0.60	−0.26	0.39
Overall		*10,088*	*71 (9)*	*58 (10)*	*29 (8)*	*1.41 (.10)*	*−0.01*	*−0.52*	*0.62*	*0.14*

Note. MOS = military occupation specialty; AFQT = Armed Forces Qualification Test (percentile score); CH = cargo helicopter; U/AH = utility/attack helicopter. The job-specific performance variable reflects the total score on a hands-on job performance test, indicating percentage of steps correctly performed (0%–100%). We report descriptive statistics for time in service (months since enlistment) before and after log-transformation. MOS component scores on four criteria based on McCloy (1994) principal component analysis. Summary statistics for each service branch and the overall sample were computed using random effects meta-analysis, pooling by MOS. Values in italics are totals and percentages for military branch subsamples and for total sample.

Data Preparation

Because time in service was positively skewed for most of the jobs, as expected (skewness ranged from −.26 to 2.08; weighted avg. = 1.00), we log-transformed this variable prior to statistical analysis. (In the results below, *time in service* refers to the log-transformed variable, unless otherwise noted.) We then screened all variables for univariate outliers within each MOS. Specifically, we replaced values on any variable exceeding ±3.5 standard deviations from the MOS sample mean with the 3.5 *SD* cutoff value. The number of univariate outliers was extremely small (30 out of 30,264, or .10% overall, and no more than .41% for any single MOS).

Corrections for Criterion Unreliability and AFQT Restriction of Range

We report observed correlations (r) and correlations corrected for statistical artifacts (r_c), including criterion unreliability and AFQT restriction of range (see Roth et al., 2017, for a discussion of range restriction). We corrected for criterion unreliability using the median internal consistency reliability values for each military branch, provided in the JPM technical report (Army: .85, Marines: .87, Air Force: .75, Navy: .81; Wigdor & Green, 1991, Volume 1, p. 121; see online Supplemental Materials, for further details on the reliability estimates used for this correction). We further corrected correlations with AFQT scores for direct restriction of range because the AFQT score is a major factor used to determine eligibility for U.S. military enlistment (Velgach & Arabian, 2023; Wigdor & Green, 1991). The unrestricted standard deviation of AFQT scores used for the range restriction correction ($SD = 28.03$) was taken from the 1980 Profile of American Youth sample (Maier, 1985, p. v). Statistical corrections for the AFQT–HOPT relationship were computed using the combined formula (No. 5) from Brown et al. (2017), which performs simultaneous corrections for direct range restriction on the predictor (i.e., AFQT scores) and criterion unreliability.[3] We were unable to correct the AFQT × MOS interaction term for statistical artifacts because there is no standard procedure for doing so. (We explored possible corrections, but none seemed viable; developing a procedure could be a useful avenue for future methodological research.) Analysis details are provided in the online Supplemental Materials; also, we present all values involved in computing corrections on the OSF https://osf.io/ap3mz/?view_only=641a709bb3394873aa70851db4b1edbb (see Oh, 2021, for a discussion of the value of this information in meta-analysis).

Meta-Analytic Approach

For all meta-analyses, we used random effects modeling, which assumes meaningful differences across effect sizes. We used restricted maximum likelihood estimation. The analyses were conducted in R (R Core Team, 2021) using the `metafor` package (Viechtbauer, 2010). Effect sizes were weighted by the inverse of their sampling error variance, such that MOSs with larger sample sizes were weighted more heavily in the model. To examine the practical and statistical significance of the heterogeneity of effects across MOS, we report the tau statistic (τ), representing an estimate of the standard deviation of the distribution of true effect sizes (once sampling error variance is taken into account), and for moderator analyses, the complementary Q statistic as an evaluation of whether the tau statistic (τ) should be interpreted (i.e., whether the null hypothesis of equal effect sizes is rejected).

Results

Table 1 presents descriptive statistics for job-specific performance, AFQT, time in service, and the four moderator variables across all MOSs. AFQT scores were highly variable across the total sample (overall $M = 54.31$, $SD = 19.99$, range = 1–99), as was time in service (overall $M = 26.10$, $SD = 14.37$, range = 1.89–160.00 months; see online Supplemental Materials).

Correlations

Table 2 presents correlations between AFQT, time in service, and job-specific performance across all MOSs. As can be seen, and as expected, AFQT generally correlated positively with job-specific performance; correlations were statistically and practically significant (median $r = .24$, median $r_c = .36$, range: −.01 to .49, range$_c$: −.02 to .69; see Figure A1). The same is true for time in service correlations with job-specific performance (median $r = .24$, median $r_c = .26$, range: −.01 to .68, range$_c$: −.01 to .72; see Figure A2). (The time in service correlations were substantially weaker for the Army, which may reflect the smaller range of values for this variable in the Army compared to the other branches.) AFQT and time in service were generally weakly correlated across MOSs (median $r = .01$, median $r_c = .02$, range: −.15 to .27, range$_c$: −.22 to .48). (See Appendix, Figures A1 and A2, for scatterplots.)

Meta-Analysis of Correlation Coefficients

Next, we meta-analyzed the observed and corrected correlations between AFQT and job-specific performance and time in service and job-specific performance. We report the results of the moderator analyses for the corrected correlations. Overall results and moderator analyses are presented in Table 3 (see Appendix, Figures A3–A5, for forest plots of effect sizes).

The meta-analytic average observed correlation between AFQT and job-specific performance was $\bar{r} = .24$, $SE = .02$, 95% CI [.20, .28], 95% credible interval, CrI [.06, .42], $p < .001$, whereas the meta-analytic corrected correlation was $\bar{r}_c = .39$, $SE = .03$, 95% CI [.34, .45], 95% CrI [.15, .63], $p < .001$. Working with things was a significant negative moderator of the relationship between AFQT and job-specific performance, indicating that jobs requiring a higher degree of manual labor tended to have a weaker relationship between AFQT scores and job-specific performance.

The meta-analytic average observed correlation between time in service and job-specific performance was $\bar{r} = .25$, $SE = .03$, 95% CI [.18, .31], 95% CrI [−.08, .57], $p < .001$, whereas the meta-analytic corrected correlation was $\bar{r}_c = .27$, $SE = .03$, 95% CI [.20, .34], 95% CrI [−.09, .63], $p < .001$. Working with things was a significant

[3] A reviewer recommended that we perform statistical corrections for direct and/or indirect range restriction on time in service, reasoning that experienced soldiers may be more likely to receive promotions (i.e., direct range restriction) and that AFQT scores and time in service were correlated for some MOS (i.e., indirect range restriction). Obviously, it would be ideal to perform such corrections, but we were unable to do so because we did not have a reference population upon which to base the corrections. Overall, the effects presented for variables involving time in service are potentially underestimates.

Table 2
Correlations of AFQT Score and Time in Service With Job-Specific Performance for 31 MOSs

			AFQT		Time in service				
MOS	Description	*N* (%)	*r* [95% CI]	r_c [95% CI]	*r* [95% CI]	r_c [95% CI]	$r_{\_diff}$	r_{c_diff}	$Z_{\_diff}$
Army	Meta-analytic estimates	*4,941 (49.0)*	***.24 [.19, .29]***	***.37 [.30, .44]***	***.06 [.01, .11]***	***.07 [.01, .12]***	*.18*	*.30*	—
11B	Infantryman	663 (6.6)	**.24 [.17, .31]**	**.33 [.23, .42]**	.01 [−.07, .08]	.01 [−.07, .09]	.23	.32	**4.34**
13B	Cannon crewman	597 (5.9)	**.12 [.04, .20]**	**.19 [.07, .31]**	**.18 [.10, .26]**	**.20 [.11, .28]**	−.06	.00	−1.01
19 E	Tank crewman	465 (4.6)	**.25 [.16, .33]**	**.35 [.23, .45]**	**.18 [.09, .27]**	**.20 [.10, .29]**	.06	.15	0.98
31C	Single-channel radio operator	346 (3.4)	**.32 [.22, .41]**	**.46 [.32, .57]**	.04 [−.07, .14]	.04 [−.07, .16]	.28	.41	**3.82**
63B	Light wheel vehicle mechanic	594 (5.9)	**.14 [.06, .22]**	**.22 [.10, .34]**	.02 [−.06, .10]	.02 [−.07, .10]	.13	.21	**2.09**
64C	Motor transport operator	646 (6.4)	**.22 [.15, .29]**	**.34 [.23, .44]**	.04 [−.03, .12]	.05 [−.04, .13]	.18	.29	**3.25**
71L	Administrative specialist	490 (4.9)	**.35 [.27, .43]**	**.52 [.41, .61]**	.01 [−.08, .09]	.01 [−.09, .10]	.34	.51	**5.27**
91A	Medical specialist	483 (4.8)	**.30 [.21, .37]**	**.49 [.36, .59]**	−.01 [−.10, .08]	−.01 [−.11, .09]	.30	.49	**4.61**
95B	Military police	657 (6.5)	**.24 [.17, .31]**	**.45 [.33, .56]**	.07 [.00, .15]	.08 [.00, .16]	.17	.37	**3.11**
Marines	Meta-analytic estimates	*3,196 (31.7)*	***.27 [.17, .37]***	***.44 [.31, .56]***	***.38 [.26, .49]***	***.40 [.28, .52]***	*−.11*	*.04*	—
031	Rifleman	940 (9.3)	**.40 [.34, .45]**	**.59 [.52, .65]**	**.33 [.27, .39]**	**.35 [.29, .41]**	.07	.23	1.61
033	Machinegunner	271 (2.7)	**.49 [.40, .58]**	**.68 [.57, .76]**	**.23 [.11, .34]**	**.24 [.12, .36]**	.27	.43	**3.55**
034	Mortarman	253 (2.5)	**.33 [.22, .44]**	**.49 [.33, .62]**	**.34 [.23, .44]**	**.36 [.24, .48]**	−.01	.13	−0.10
035	Assaultman	277 (2.7)	**.37 [.26, .47]**	**.56 [.42, .67]**	**.20 [.09, .31]**	**.22 [.09, .34]**	.17	.34	**2.15**
3521	Organizational automotive mechanic	907 (9.0)	**.25 [.19, .31]**	**.44 [.34, .53]**	**.20 [.14, .26]**	**.21 [.15, .28]**	.05	.22	1.08
6112	Helicopter mechanic CH-46	152 (1.5)	.13 [−.03, .28]	.24 [−.06, .49]	**.63 [.52, .71]**	**.67 [.55, .76]**	−.50	−.43	**−4.92**
6113	Helicopter mechanic CH-53	93 (0.9)	.12 [−.09, .32]	.21 [−.15, .51]	**.68 [.55, .77]**	**.72 [.59, .83]**	−.56	−.52	**−4.43**
6114	Helicopter mechanic U/AH-1	190 (1.9)	**.21 [.07, .34]**	**.36 [.13, .55]**	**.41 [.28, .52]**	**.43 [.30, .55]**	−.19	−.08	**−2.01**
6115	Helicopter mechanic CH-53 E	113 (1.1)	−.01 [−.20, .17]	−.02 [−.33, .30]	**.38 [.21, .53]**	**.41 [.23, .57]**	−.40	−.43	**−2.92**
Air Force	Meta-analytic estimates	*1,244 (12.3)*	***.21 [.14, .28]***	***.38 [.26, .51]***	***.32 [.25, .38]***	***.37 [.29, .44]***	*−.11*	*.01*	—
112	Aircrew life support specialist	166 (1.6)	.10 [−.05, .25]	.19 [−.10, .44]	**.35 [.21, .48]**	**.41 [.24, .55]**	−.25	−.22	**−2.76**
272	Air traffic control operator	171 (1.7)	.10 [−.05, .25]	.22 [−.10, .49]	**.26 [.12, .40]**	**.30 [.14, .46]**	−.16	−.08	−1.47
324	Precision measuring equipment specialist	124 (1.2)	**.30 [.13, .45]**	**.62 [.31, .80]**	**.37 [.20, .51]**	**.42 [.23, .59]**	−.07	.20	−0.63
328	Avionic communications specialist	83 (0.8)	**.33 [.12, .51]**	**.69 [.31, .86]**	**.30 [.09, .48]**	**.34 [.10, .56]**	.03	.35	0.20
423	Aerospace ground equipment specialist	216 (2.1)	**.18 [.05, .31]**	**.33 [.09, .53]**	**.24 [.11, .36]**	**.28 [.13, .42]**	−.06	.05	−0.69
426	Jet engine mechanic	188 (1.9)	.10 [−.04, .24]	.18 [−.07, .40]	**.18 [.04, .32]**	**.21 [.05, .37]**	−.08	−.03	−0.71
492	Information systems radio operator	120 (1.2)	**.32 [.15, .47]**	**.51 [.25, .69]**	**.43 [.27, .56]**	**.49 [.31, .65]**	−.11	.01	−0.99
732	Personnel specialist	176 (1.7)	**.29 [.15, .42]**	**.50 [.27, .67]**	**.40 [.27, .52]**	**.47 [.31, .60]**	−.12	.04	−1.20
Navy	Meta-analytic estimates	*707 (7.00)*	***.19 [.12, .26]***	***.31 [.19, .43]***	***.25 [.18, .31]***	***.27 [.20, .35]***	*−.06*	*.04*	—
EM	Electrician's mate	80 (0.8)	**.27 [.05, .46]**	**.40 [.08, .64]**	**.36 [.15, .54]**	**.40 [.17, .60]**	−.09	.00	−0.64
ET	Electronics technician	136 (1.3)	.13 [−.04, .29]	.27 [−.09, .56]	**.25 [.08, .40]**	**.28 [.09, .44]**	−.12	.00	−1.18
GSM	Gasoline turbine mechanic	89 (0.9)	.13 [−.08, .33]	.20 [−.13, .48]	**.23 [.02, .42]**	**.26 [.03, .47]**	−.10	−.06	−0.66
MM	Machinist's mate	178 (1.8)	**.20 [.05, .34]**	**.27 [.07, .44]**	**.20 [.05, .34]**	**.22 [.06, .37]**	.00	.05	0.00
RM	Radioman	224 (2.2)	**.21 [.09, .34]**	**.42 [.18, .60]**	**.24 [.11, .36]**	**.26 [.12, .40]**	−.02	.15	−0.27
—	Overall meta-analytic estimates	*10,088 (100)*	***.24 [.20, .28]***	***.39 [.34, .45]***	***.25 [.18, .31]***	***.27 [.20, .34]***	*−.01*	*.12*	—

Note. MOS = military occupation specialty; AFQT = Armed Forces Qualification Test (percentile score); CH = cargo helicopter; U/AH = utility/attack helicopter; HOPT = hands-on job performance test. The job-specific performance variable reflects total score on a HOPT, indicating percentage of steps correctly performed (0%–100%). We report descriptive statistics for time in service (months since enlistment) before and after log-transformation. 95% CI = 95% confidence interval. r_c, correlation after correction for job-specific performance unreliability and AFQT restriction of range. $r_{\_diff}$, raw difference between observed correlations. r_{c_diff}, raw difference between corrected correlations. $Z_{\_diff}$, Steiger's test of difference between dependent observed correlations. Boldface indicates $p < .05$. Values in italics are totals and percentages for military branch subsamples and for total sample.

Table 3

Meta-Analysis of Observed and Corrected Correlations Between Job-Specific Performance and AFQT Score and Time in Service

Measure	Meta-analytic average correlation	τ	Moderators				
			Military branch	Cognitive complexity	Working with things	Unpleasant working conditions	Fine motor control
AFQT (observed)	$\bar{r} = .24$, $SE = .02$, $p < .001$ 95% CI [.20, .28] 95% CrI [.06, .42]	$\tau = .09$ 95% CI [.06, .13]	—	—	—	—	—
AFQT (corrected)	$\bar{r}_c = .39$, $SE = .03$, $p < .001$ 95% CI [.34, .45] 95% CrI [.15, .63]	$\tau = .12$ 95% CI [.08, .18]	Q_M (3) = 3.33 $p = .34$	Q_M (1) = 1.98 $b = -.03$ $SE = .02$ $p = .16$	**Q_M (1) = 6.35** **$b = -.08$** **$SE = .03$** **$p = .01$**	Q_M (1) = 0.26 $b = .01$ $SE = .03$ $p = .61$	Q_M (1) = 0.14 $b = .03$ $SE = .07$ $p = .71$
Time in service (observed)	$\bar{r} = .25$, $SE = .03$, $p < .001$ 95% CI [.18, .31] 95% CrI [−.08, .57]	$\tau = .16$ 95% CI [.12, .22]	—	—	—	—	—
Time in service (corrected)	$\bar{r}_c = .27$, $SE = .03$, $p < .001$ 95% CI [.20, .34] 95% CrI [−.09, .63]	$\tau = .18$ 95% CI [.14, .24]	**Q_M (3) = 39.06** **$p < .001$**	Q_M (1) = 1.57, $b = .04$ $SE = .03$ $p = .21$	**Q_M (1) = 8.94** **$b = .10$** **$SE = .03$** **$p = .003$**	Q_M (1) = 0.18 $b = -.01$ $SE = .03$ $p = .68$	Q_M (1) = 1.18 $b = -.09$ $SE = .08$ $p = .28$

Note. $k = 31$. $N = 10{,}088$. Boldface indicates that the moderator analysis was statistically significant. Dashes indicate that moderator analyses were only conducted on the corrected correlations. AFQT = Armed Forces Qualification Test; SE = standard error; CI = confidence interval; CrI = credible interval.

positive moderator of the relationship between time in service and job-specific performance, indicating that jobs requiring more manual labor tended to have a stronger relationship between time in service and job-specific performance.

Regression Analyses

Our major research question was whether AFQT score (an index of g) and time in service (an index of job experience) would interact in predicting job-specific performance. For each of the 31 MOSs, we performed a hierarchical regression analysis, with AFQT and time in service entered in Step 1 as the main effects and the AFQT × Time in Service cross-product interaction term added in Step 2. Regarding the interaction term, the effect of AFQT on job-specific performance could decrease (the convergence hypothesis), increase (the divergence hypothesis), or remain the same (the noninteractive hypothesis) as time in service increases.

Regression results are displayed in Table 4. The main effects of both AFQT and time in service were positive for all MOSs and statistically significant for most MOSs (26 of 31 for AFQT, 24 of 31 for time in service). High levels of both g and time in service were associated with higher levels of job-specific performance (βs ranged from .01 to .48 for AFQT and from .01 to .67 for time in service). The main effects accounted for a considerable proportion of the variance in job-specific performance (R ranged from .14 to .69, $Mdn = .33$). However, the median AFQT × Time in Service interaction was .00. Moreover, statistical significance was found for only two of the 31 MOS interactions (Army administrative specialist: $\beta = -.11$, $t = -2.55$, $p = .011$, $r_{\text{part}} = -.11$; Marines mortarman, $\beta = -.11$, $t = -2.05$, $p = .042$, $r_{\text{part}} = -.11$), and the values were near-zero and would not be statistically significant after performing a Bonferroni correction for multiple comparisons ($\alpha = .05/31 = .0016$).

In summary, these large-scale results provide essentially no evidence that time in service moderated the effect of AFQT score on job-specific performance. Instead, the results support the noninteractive (main effects) hypothesis.[4]

Meta-Analysis of Regression Coefficients

Next, we (a) computed the average effects across the 31 MOSs for AFQT, time in service, and the AFQT × Time in Service interaction on job-specific performance, and (b) performed moderator analyses considering the military branch and the four aforementioned job characteristics: cognitive complexity, working with things, unpleasant working conditions, and fine motor control. For each regression analysis (MOS) in Table 4, we entered into the meta-analysis the regression coefficient (as the effect) and its corresponding squared standard error (as the inverse weight of the effect). We used random effects modeling and restricted maximum likelihood estimation. The analyses were conducted in R (R Core Team, 2021) using the `metafor` package (Viechtbauer, 2010). We report the meta-analytically weighted average effect for the main effects of AFQT and time in service; then we meta-analyze the standardized regression coefficients for the AFQT × Time in Service interaction after taking the

[4] See Table A2 for hierarchical regression analyses showing incremental validity of AFQT and Time in Service for Job-Specific Performance Test Scores computed using corrected correlations.

Table 4
Regression Analyses Predicting Job-Specific Performance for 31 MOSs

			Model Step 1		AFQT			Time in service			Model Step 2				AFQT × TIS		
MOS	Description	N (%)	R	F	β	SE	r_{part}	β	SE	r_{part}	R	F	ΔR	F_{ch}	β	SE	r_{part}
Army																	
11B	Infantryman	663 (6.6)	**.24**	**20.54**	**.24**	**.04**	**.24**	.01	.04	.01	**.25**	**14.10**	.00	1.21	−.04	.04	−.04
13B	Cannon crewman	597 (5.9)	**.22**	**14.55**	**.12**	**.04**	**.12**	**.18**	**.04**	**.18**	**.22**	**9.77**	.00	.24	.02	.04	.02
19E	Tank crewman	465 (4.6)	**.31**	**23.86**	**.24**	**.04**	**.24**	**.18**	**.04**	**.18**	**.31**	**16.53**	.01	1.80	−.06	.04	−.06
31C	Single-channel radio operator	346 (3.4)	**.32**	**19.31**	**.32**	**.05**	**.32**	.03	.05	.03	**.32**	**12.84**	.00	.01	.00	.05	.00
63B	Light wheel vehicle mechanic	594 (5.9)	**.14**	**6.33**	**.15**	**.04**	**.15**	.02	.04	.02	**.15**	**4.61**	.01	1.17	−.04	.04	−.04
64C	Motor transport operator	646 (6.4)	**.22**	**17.08**	**.22**	**.04**	**.22**	.04	.04	.04	**.23**	**11.57**	.00	.58	.03	.04	.03
71L	Administrative specialist	490 (4.9)	**.35**	**34.61**	**.37**	**.04**	**.36**	.05	.04	.05	**.37**	**25.50**	**.02**	**6.51**	**−.11**	**.04**	**−.11**
91A	Medical specialist	483 (4.8)	**.30**	**22.89**	**.30**	**.04**	**.30**	.02	.04	.02	**.30**	**15.47**	.00	.65	−.04	.04	−.04
95B	Military police	657 (6.5)	**.25**	**22.07**	**.24**	**.04**	**.24**	.07	.04	.07	**.25**	**14.72**	.00	.09	.01	.04	.01
Marines																	
031	Rifleman	940 (9.3)	**.50**	**159.28**	**.38**	**.03**	**.38**	**.31**	**.03**	**.31**	**.50**	**106.33**	.00	.59	−.02	.03	−.02
033	Machinegunner	271 (2.7)	**.53**	**52.85**	**.48**	**.05**	**.47**	**.20**	**.05**	**.20**	**.53**	**35.11**	.00	.02	−.01	.05	−.01
034	Mortarman	253 (2.5)	**.46**	**34.02**	**.31**	**.06**	**.31**	**.32**	**.06**	**.32**	**.48**	**24.36**	**.01**	**4.19**	**−.11**	**.06**	**−.11**
035	Assaultman	277 (2.7)	**.41**	**28.04**	**.36**	**.06**	**.36**	**.18**	**.06**	**.18**	**.41**	**18.65**	.00	.06	−.01	.06	−.01
3521	Organizational automotive mechanic	907 (9.0)	**.33**	**55.32**	**.26**	**.03**	**.26**	**.21**	**.03**	**.21**	**.33**	**37.34**	.00	1.33	.04	.03	.04
6112	Helicopter mechanic CH-46	152 (1.5)	**.65**	**53.18**	**.15**	**.06**	**.15**	**.64**	**.06**	**.64**	**.65**	**36.09**	.00	1.53	.08	.06	.08
6113	Helicopter mechanic CH-53	93 (0.9)	**.69**	**41.70**	**.16**	**.08**	**.16**	**.67**	**.08**	**.66**	**.70**	**28.20**	.00	1.11	−.08	.08	−.08
6114	Helicopter mechanic U/AH-1	190 (1.9)	**.46**	**25.07**	**.22**	**.06**	**.22**	**.40**	**.07**	**.40**	**.47**	**17.50**	.01	2.07	−.09	.07	−.09
6115	Helicopter mechanic CH-53 E	113 (1.1)	**.38**	**9.54**	.03	.08	.03	**.38**	**.08**	**.37**	**.40**	**6.90**	.01	1.52	−.11	.08	−.11
Air Force																	
112	Aircrew life support specialist	166 (1.6)	**.35**	**11.56**	.01	.08	.01	**.34**	**.08**	**.33**	**.36**	**8.11**	.01	1.19	−.08	.07	−.08
272	Air traffic control operator	171 (1.7)	**.29**	**7.58**	.12	.07	.12	**.27**	**.07**	**.27**	**.29**	**5.02**	.00	.00	.00	.07	.00
324	Precision measuring equipment specialist	124 (1.2)	**.43**	**13.76**	**.24**	**.08**	**.23**	**.32**	**.08**	**.31**	**.43**	**9.16**	.00	.15	−.03	.08	−.03
328	Avionic communications specialist	83 (0.8)	**.44**	**9.79**	**.32**	**.10**	**.32**	**.30**	**.10**	**.30**	**.46**	**6.97**	.01	1.26	−.11	.10	−.11
423	Aerospace ground equipment specialist	216 (2.1)	**.29**	**10.01**	**.16**	**.07**	**.16**	**.23**	**.07**	**.23**	**.29**	**6.65**	.00	.04	.01	.07	.01
426	Jet engine mechanic	188 (1.9)	**.23**	**5.02**	.14	.07	.13	**.20**	**.07**	**.20**	**.23**	**3.33**	.00	.01	.01	.07	.01
492	Information systems radio operator	120 (1.2)	**.50**	**19.75**	**.26**	**.08**	**.26**	**.39**	**.08**	**.38**	**.50**	**13.09**	.00	.09	−.02	.08	−.02
732	Personnel specialist	176 (1.7)	**.49**	**27.28**	**.28**	**.07**	**.28**	**.40**	**.07**	**.40**	**.49**	**18.13**	.00	.10	−.02	.07	−.02
Navy																	
EM	Electrician's mate	80 (0.8)	**.44**	**9.07**	**.36**	**.12**	**.29**	**.33**	**.10**	**.33**	**.47**	**7.00**	.03	2.51	−.20	.12	−.16
ET	Electronics technician	136 (1.3)	**.26**	**4.65**	.06	.09	.06	**.23**	**.09**	**.22**	**.27**	**3.39**	.01	.87	.08	.08	.08
GSM	Gasoline turbine mechanic	89 (0.9)	**.28**	**3.67**	.15	.10	.15	**.23**	**.11**	**.23**	**.30**	**2.75**	.02	.92	.10	.11	.10
MM	Machinist's mate	178 (1.8)	**.29**	**8.23**	**.23**	**.07**	**.22**	**.23**	**.07**	**.22**	**.30**	**5.62**	.00	.46	−.05	.08	−.05
RM	Radioman	224 (2.2)	**.30**	**11.26**	**.20**	**.07**	**.20**	**.21**	**.06**	**.21**	**.31**	**7.81**	.01	.92	−.06	.06	−.06

Note. MOS = military occupation specialty; AFQT = Armed Forces Qualification Test (percentile score); TIS = time in service (log months); CH = cargo helicopter; U/AH = utility/attack helicopter. The job-specific performance variable is total score on a hands-on job performance test, indicating percentage of steps correctly performed (0–100). Parameter estimates are from the full model with all variables entered into the regression equation. F_{ch} = F statistic for change in fit from Step 1 to Step 2. Boldface indicates $p < .05$. SE = standard error.

main effects into account. Throughout, we report results of the moderator analyses.

AFQT

The meta-analytic average effect of AFQT on job-specific performance was $\beta = .23$, $SE = .02$, 95% CI [.20, .27], 95% CrI [.06, .41], $p < .001$, $\tau = .09$, 95% CI [.06, .13]; β was positive for all 31 MOSs (Figure A3). For every 1 standard deviation increase in AFQT scores, job-specific performance increased by approximately one quarter of 1 standard deviation. Military branch was not a statistically significant moderator, Q_M (3) = 3.95, $p = .27$. Cognitive complexity significantly moderated the relationship between AFQT and job-specific performance, Q_M (1) = 4.35, $b = -.03$, $SE = .02$, $p = .037$; more cognitively complex MOSs had a weaker relationship between AFQT scores and job-specific performance. Working with things also significantly moderated the relationship between AFQT and job-specific performance: Q_M (1) = 4.47, $b = -.04$, $SE = .02$, $p = .035$; MOSs that involved more manual labor had a weaker relationship between AFQT scores and job-specific performance. The other moderators were not statistically significant. Unpleasant working conditions: Q_M (1) = 2.03, $b = .03$, $SE = .02$, $p = .15$. Fine motor control: Q_M (1) = 0.36, $b = -.03$, $SE = .05$, $p = .55$.

Time in Service

The meta-analytic average effect of time in service on job-specific performance was $\beta = .24$, $SE = .03$, 95% CI [.18, .30], 95% CrI [−.07, .54], $p < .001$, $\tau = .15$, 95% CI [.11, .21]; β was positive for all 31 MOSs (Figure A4). For every 1 standard deviation increase in time in service, job-specific performance increased by approximately one quarter of 1 standard deviation. The effect of time in service on job-specific performance was significantly moderated by military branch, Q_M (3) = 38.34, $p < .001$. Average time in service effects were: Marines, $\beta = .36$, $SE = .06$, 95% CI [.25, .48], 95% CrI [.02, .72], $p < .001$; Air Force, $\beta = .30$, $SE = .03$, 95% CI [.25, .36], 95% CrI [.24,

.37], $p < .001$; Navy, β = .24, SE = .04, 95% CI [.17, .31], 95% CrI [.17, .31], $p < .001$; and Army, β = .07, SE = .02, 95% CI [.02, .11], 95% CrI [−.05, .18], p = .003. Thus, the effect of time in service on job-specific performance was practically and significantly lower for Army MOSs than for MOSs in the other branches.

The effect of time in service was not moderated by cognitive complexity, Q_M (1) = 1.78, b = .03, SE = .03, p = .18, but it was so for working with things, Q_M (1) = 9.24, b = .09, SE = .03, p = .002; the relationship between time in service and job-specific performance was stronger for jobs that involved more manual labor. The moderating effects of unpleasant working conditions and fine motor control were not statistically significant, Q_M (1) = 0.07, b = −.01, SE = .03, p = .79, and Q_M (1) = 1.45, b = −.09, SE = .07, p = .23, respectively.

AFQT × Time in Service

Although very small in magnitude, the meta-analytic average effect was statistically significant, β = −.02, SE = .01, 95% CI [−.04, −.00], 95% CrI [−.05, .01], p = .027; β was negative (the direction predicted by the convergence hypothesis, Figure A5). The magnitude of the interaction effect was similar across MOSs, as indicated by a low degree of between-study variability: τ = .01, 95% CI [.00, .04]. Consistent with this result for no overall interaction effect or heterogeneity in the effect, the AFQT × Time in Service interaction effect was not moderated to a statistically significant degree by military branch (p = .99) or by any of the four job characteristics (all ps > .31).

Discussion

Measures of g positively, meaningfully, and consistently predict job-specific performance across a wide range of jobs (Schmidt & Hunter, 2004). However, the question of whether job experience moderates the relationship between g and job-specific performance has remained open. Using data from the JPM project, we tested three competing hypotheses concerning this question: the convergence hypothesis, the divergence hypothesis, and the noninteractive hypothesis (Schmidt et al., 1988). This study was the first to include all the following features: (a) a large sample size (N = 10,088) to provide sufficient statistical power to detect small but still theoretically and practically meaningful effects; (b) a diverse range of jobs (k = 31 MOSs) across which the criterion measures of task performance differed in ways appropriate to the job but were all behavioral measures scored by trained raters; (c) direct tests for main and interactive effects of cognitive ability and job experience on job-specific performance, for each job and across jobs; and (d) use of meta-analysis to integrate the results.

As expected, the main effects of both AFQT and time in service were positive for all 31 MOSs, and statistically significant for most of them. There was also some evidence for moderation of these relationships. For jobs high on working with things, the positive relationship between AFQT and job-specific performance was weaker, whereas the positive relationship between time and service and job-specific performance was stronger. In other words, AFQT tended to be a weaker (less positive) predictor of job-specific in MOSs requiring a high level of manual labor, whereas time in service tended to be a stronger (more positive) predictor of job-specific performance in these MOSs. One possible interpretation of this pattern of results is that jobs that require a higher degree of working with things place lower demands on cognitive resources (as opposed to physical resources), accounting for the weaker effect of AFQT, and are also more automatable through experience, accounting for the stronger effect of time in service.

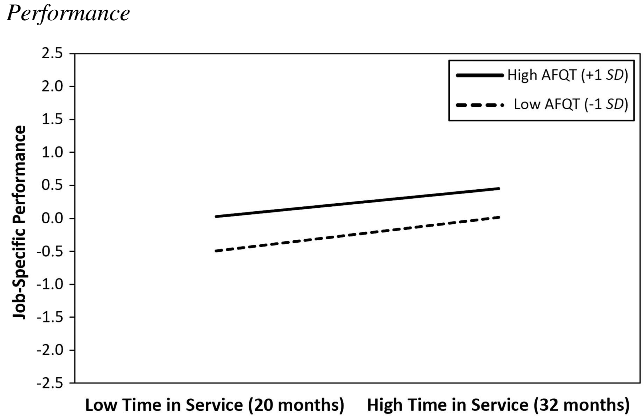
More central to our research question, the AFQT × Time in Service interaction was negligible at both the individual MOS level and the meta-analytic level. The meta-analytic average regression coefficient for this interaction was β = −.02, 95% CI [−.04, .00]. Conceptually, this means that increases in time in service resulted in negligible decreases in the predictive validity of AFQT for job-specific performance. Statistically, as illustrated in Figure 2, this means that for every 1 standard deviation increase in time in service above the sample mean, the predictive validity of AFQT for job-specific performance decreased by only .02.

The take-home message of this study is that g remains a statistically and practically significant predictor of job-specific performance, even in employees with high levels of job experience. Reeve and Bonaccio (2011) reached a similar conclusion in a narrative review, 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) and that the hypothesis of diminishing validity of g "appears to be a myth" (p. 268). The present study bolsters this conclusion and, we believe, provides the strongest evidence against the still-popular convergence hypothesis (see Ericsson, 2014; Ericsson & Pool, 2016; Ward et al., 2017).

Sackett et al. (2022) recently argued that meta-analytic estimates of the validity of g for job performance are typically inflated due to overcorrection for range restriction in g. Accordingly, Sackett et al. advised that "if one is not confident in the basis for a range restriction correction, it is better to forego a correction than to use a value that results in an overestimate" (pp. 60–61). We were, in fact, confident in the basis for the range restriction correction we applied to the correlations between AFQT and job-specific performance. As

Figure 2

Plot of the Negligible Meta-Analytic Average Interaction Effect of AFQT (Percentile Score) and Time in Service on Job-Specific Performance



Note. Unit of measurement for job-specific performance is the z score of the hands-on job performance score. Predicted slopes are at ±1 SD for AFQT (+1 SD = 68.19, −1 SD = 47.91) and for log time in service (+1 SD = 32 months in original units, −1 SD = 20 months in original units) based on the meta-analytic results. AFQT = Armed Forces Qualification Test.

already mentioned, the AFQT is a major factor in determining whether an individual is eligible to enlistment in the U.S. military. This presumably leads to direct range restriction in AFQT scores of enlistees. Thus, we corrected for direct range restriction using the AFQT *SD* from a nationally representative sample of American young adults (18–23 years). Notably, the meta-analytic average correlation between AFQT and job-specific performance we observed in this study ($r_c = .39$) is similar to Sackett et al.'s estimate of .31 for the validity of *g* as a predictor of job performance, and considerably lower than the meta-analytic average reported by others (e.g., .51; Hunter & Hunter, 1984; Schmidt & Hunter, 2004). This finding lends credence to Sackett et al.'s (2022) argument that validity estimates for *g* have been overestimated in the past, even if they are still of a statistically and practically meaningful magnitude as predictors of job-specific performance.

Why does *g* remain important as a predictor of job-specific performance even after extensive job experience? It could be that most complex job tasks include not only *consistent* demands, where the rules are either well understood from the outset or become well understood as skill is acquired, but also retain *variable* demands, where the rules and situations change over time. Automaticity may emerge in consistent job components but not in variable components. *g* may remain important for these latter components. Also, underlying this overall effect, there may well be *components* of certain jobs where the effect of *g* on job-specific performance might diminish. For example, Hambrick et al. (2012) found evidence for an underadditive interaction between visuospatial ability and geological knowledge on geologists' success in a bedrock mapping task (for another example, see Sohn & Doane, 2003).

Two other findings are worth briefly mentioning. First, mean time in service correlated significantly ($r = .49$) with mean AFQT score across MOSs, suggesting that higher ability personnel stayed in the military longer. Second, consistent with previous findings (e.g., Avolio & Waldman, 1990), mean AFQT score and cognitive complexity across MOSs were positively correlated ($r = .49$; see Table A1), suggesting that higher ability individuals tended to be selected for more cognitively complex jobs (via personnel selection and/or self-selection processes). Additionally, there may have been greater restriction of range in cognitively complex jobs than in less cognitively complex jobs. Across MOSs, the cognitive complexity of the MOS was strongly and negatively related to the standard deviation of the MOS's mean AFQT score ($r = -.44$), indicating that variability in AFQT scores was lower for jobs with high cognitive complexity than for jobs with low cognitive complexity. By contrast, when conducting moderator analyses on the *restriction-of-range-corrected correlation* between AFQT scores and job performance, cognitive complexity was not a statistically significant moderator. This pattern of results speaks to the importance of accounting for statistical artifacts where possible.

Limitations of Study

We note several limitations of this study. First, the study was cross-sectional. As Schmidt et al. (1988) noted, the ideal way to investigate the joint effects of cognitive ability and job experience on job performance is to conduct longitudinal studies. In particular, they cautioned that changes in selection standards and/or selective attrition of participants can lead to the false appearance of interactions supporting either the convergence or divergence hypothesis. We agree that longitudinal studies are ideal for our research question, but note that we found no evidence for either interaction hypothesis. Nonetheless, the cross-sectional nature of the present study is a prominent limitation.

Second, given that some personnel might have switched MOS from their original training, time in service is not a perfect measure of job experience. However, any MOS switching (reclassification) was likely to have been into related MOS where relevant job experience would transfer (R. McCloy, personal communication, May 5, 2020). Third, in terms of the widely accepted Cattell–Horn–Carroll model of intelligence (Cattell, 1943; Horn & Cattell, 1967), the AFQT score is weighted toward *crystallized* abilities (Gc) reflecting knowledge/skill acquired through experience, as opposed to *fluid* reasoning abilities (Gf; Roberts et al., 2000). Gc has been found to be a stronger predictor of job performance than Gf (Postelthwaite, 2011), but Gf may be more predictive of the more variable aspects of work that require abstract reasoning, thereby adding to the prediction of job-specific performance.

Fourth, participants in the JPM study knew they were being evaluated in the HOPTs. Thus, their work motivation reflected maximal and not typical levels of motivation, which can affect both mean performance and correlations with performance (Sackett et al., 1988). Fifth, there was a restricted range of MOS job complexity, as job complexity ranged from low to medium. Thus, although our results are wide-ranging across 31 MOSs, it is an open question whether they would be similar for jobs at higher levels of complexity. Finally, the JPM data set is nearly 40 years old, and some of the jobs in the study have no doubt changed (Oswald et al., 2019). However, we still suspect strong generalizability of findings to the world of work, given the diverse range of jobs in our data set. Moreover, we know of no evidence to indicate *g* has become less (or more) important as a predictor of job-specific performance over historical time.

A Broader Direction for Future Research

We believe that future research should investigate the joint effects of cognitive ability and job experience on job-specific performance at a more cognitive or mechanistic level. One way to think about the convergence hypothesis is in terms of *circumvention of processing limitations* (Hambrick & Meinz, 2011; Salthouse, 1991). The idea is that the acquisition of domain-specific knowledge/skills enables performers to bypass reliance on capacity-limited aspects of cognition such as working memory. For example, not unlike how the chess master can often look at a chessboard and quickly determine the best move (Chase & Simon, 1973), a skilled mechanic may be able to diagnose a problem with an automobile through automatic recognition of the car's symptoms. Specification of mechanisms underlying effects of cognitive ability, personality, interests, and other relatively stable worker characteristics—as well as a deep understanding of occupational characteristics that place specific demands on worker characteristics (e.g., those from the Occupational Information Network Content Model; Peterson et al., 2001)—will help applied psychologists adopt a more integrated and more powerful "whole person" and "whole job" approach to the predictions and processes underlying job-specific performance.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of*

Experimental Psychology: General, *117*(3), 288–318. https://doi.org/10.1037/0096-3445.117.3.288

Avolio, B. J., & Waldman, D. A. (1990). An examination of age and cognitive test performance across job complexity and occupational types. *Journal of Applied Psychology*, *75*(1), 43–50. https://doi.org/10.1037/0021-9010.75.1.43

Brown, R. D., Oswald, F. L., & Converse, P. D. (2017). Estimating operational validity under incidental range restriction: Some important but neglected issues. *Practical Assessment, Research & Evaluation*, *22*, Article 6. https://doi.org/10.7275/e7gh-0785

Campbell, J. P., & Wiernik, B. M. (2015). The modeling and assessment of work performance. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 47–74. https://doi.org/10.1146/annurev-orgpsych-032414-111427

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*(3), 153–193. https://doi.org/10.1037/h0059973

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2

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. https://doi.org/10.1016/j.intell.2013.12.001

Ericsson, K. A. (2018). The differential influence of experience, practice, and deliberate practice on the development of superior individual performance of experts. In K. A. Ericsson, R. R. Hoffmann, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed., pp. 745–769). Cambridge University Press. https://doi.org/10.1017/9781316480748.038

Ericsson, K. A., & Pool, R. (2016). *Peak: Secrets from the new science of expertise*. Houghton Mifflin Harcourt.

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*(1), 60–79. https://doi.org/10.1037/0021-9010.86.1.60

Fleishman, E. A., & Rich, S. (1963). Role of kinesthetic and spatialvisual abilities in perceptual-motor learning. *Journal of Experimental Psychology*, *66*(1), 6–11. https://doi.org/10.1037/h0046677

Gade, P. A., & Dudley, N. M. (2004). *Sixty years of U.S. Army selection and classification test development*. U.S. Army Research Institute for the Behavioral and Social Sciences. https://www.researchgate.net/profile/Paul-Gade/publication/228457064_Sixty_Years_of_US_Army_Selection_and_Classification_Test_Development/links/55eeee5108ae199d47bfe5dc/Sixty-Years-of-US-Army-Selection-and-Classification-Test-Development.pdf

Hambrick, D. Z., Burgoyne, A. P., & Oswald, F. L. (2019). Domain-general models of expertise: The role of cognitive ability. In P. Ward, J. M. Schraagen, J. Gore, & E. Roth (Eds.), *The Oxford handbook of expertise: Research and application* (pp. 56–84). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198795872.013.3

Hambrick, D. Z., Libarkin, J. C., Petcovic, H. L., Baker, K. M., Elkins, J., Callahan, C. N., Turner, S. P., Rench, T. A., & 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*(3), 397–403. https://doi.org/10.1037/a0025927

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*(5), 275–279. https://doi.org/10.1177/0963721411422061

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*(2), 107–129. https://doi.org/10.1016/0001-6918(67)90011-X

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, *96*(1), 72–98. https://doi.org/10.1037/0033-2909.96.1.72

Johnson, J. W. (2001). The relative importance of task and contextual performance dimensions to supervisor judgments of overall performance. *Journal of Applied Psychology*, *86*(5), 984–996. https://doi.org/10.1037/0021-9010.86.5.984

Keil, C. T., & Cortina, J. M. (2001). Degradation of validity over time: A test and extension of Ackerman's model. *Psychological Bulletin*, *127*(5), 673–697. https://doi.org/10.1037/0033-2909.127.5.673

Lance, C. E., Hedge, J. W., & Alley, W. E. (1989). Joint relationships of task proficiency with aptitude, experience, and task difficulty: A cross-level, interactional study. *Human Performance*, *2*(4), 249–272. https://doi.org/10.1207/s15327043hup0204_2

Maier, M. H. (1985). *Effects of truncating a reference population on correction of validity coefficients for range restriction*. Center for Naval Analyses, Marine Corps Operations Analysis Group. https://doi.org/10.21236/ADB103741

Mayberry, P. W., & Carey, N. B. (1997). The effect of aptitude and experience on mechanical job performance. *Educational and Psychological Measurement*, *57*(1), 131–149. https://doi.org/10.1177/0013164497057001010

McCloy, R. A. (1994, January). Predicting job performance scores without performance data. In B. F. Green Jr., & A. S. Mavor (Eds.), *Modeling cost and performance for military enlistment: Report of a workshop* (pp. 61–100). National Academy Press. https://doi.org/10.17226/2344

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.

Oh, I.-S. (2021). Challenges and pitfalls of Cinderella-izing descriptive statistics in Table 1. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *14*(4), 521–523. https://doi.org/10.1017/iop.2021.105

Oswald, F. L., Behrend, T. S., & Foster, L. L. (Eds.). (2019). *Workforce readiness and the future of work*. Taylor & Francis. https://doi.org/10.4324/9781351210485

Palmer, P., Hartke, D. D., Ree, M. J., Welsh, J. R., & Valentine, L. D., (1988). *Armed Services Vocational Aptitude Battery (ASVAB): Alternate forms reliability (Forms 8, 9, 10, and 11)* [Database record]. Accession number: ADA191658. https://doi.org/10.1037/e532982009-001

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Levin, K. Y., Campion, M. A., Mayfield, M. S., Morgeson, F. P., Pearlman, K., Gowing, M. K., Lancaster, A. R., Silver, M. B., & Dye, D. M. (2001). Understanding work using the occupational information Network (O* NET): Implications for practice and research. *Personnel Psychology*, *54*(2), 451–492. https://doi.org/10.1111/j.1744-6570.2001.tb00100.x

Postelthwaite, B. E. (2011). *Fluid ability, crystallized ability, and performance across multiple domains: A meta-analysis* [Unpublished doctoral dissertation]. The University of Iowa.

Quiñones, M. A., Ford, J. K., & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology*, *48*(4), 887–910. https://doi.org/10.1111/j.1744-6570.1995.tb01785.x

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

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*(5), 255–272. https://doi.org/10.1016/j.intell.2011.06.009

Roberts, R. D., Goff, G. N., Anjoul, F., Kyllonen, P. C., Pallier, G., & Stankov, L. (2000). The Armed Services Vocational Aptitude Battery (ASVAB): Little more than acculturated learning (Gc)!? *Learning and Individual Differences*, *12*(1), 81–103. https://doi.org/10.1016/S1041-6080(00)00035-2

Roth, P. L., Le, H., Oh, I.-S., Van Iddekinge, C. H., & Robbins, S. B. (2017). Who r u? On the (in)accuracy of incumbent-based estimates of range restriction in criterion-related and differential validity research. *Journal of Applied Psychology*, *102*(5), 802–828. https://doi.org/10.1037/apl0000193

Rothe, H. F. (1946). Output rates among butter wrappers; work curves and their stability. *Journal of Applied Psychology*, *30*(3), 199–211. https://doi.org/10.1037/h0062587

Sackett, P. R., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximum job performance. *Journal of Applied Psychology*, *73*(3), 482–486. https://doi.org/10.1037/0021-9010.73.3.482

Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, *107*(11), 2040–2068. https://doi.org/10.1037/apl0000994

Salthouse, T. A. (1991). Expertise as circumvention of normal processing limitations. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 286–300). Cambridge University Press.

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*(1), 162–173. https://doi.org/10.1037/0022-3514.86.1.162

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*(1), 46–57. https://doi.org/10.1037/0021-9010.73.1.46

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*(3), 458–466. https://doi.org/10.3758/BF03194403

Velgach, S., & Arabian, J. M. (2023). Commentary on the criterion problem in the US Department of Defense: Policy and operational considerations. *Military Psychology*, *35*(4), 364–371. https://doi.org/10.1080/08995605.2023.2218787

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viswesvaran, C., & Ones, D. S. (2002). Agreements and disagreements on the role of general mental ability (GMA) in industrial, work, and organizational psychology. *Human Performance*, *15*(1–2), 211–231. https://doi.org/10.1080/08959285.2002.9668092

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). Routledge. https://doi.org/10.4324/9781315668017-3

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

Wise, L. L. (1994). Setting performance goals for the DoD linkage model. In B. F. Green & A. S. Mavor (Eds.), *Modeling cost and performance for military enlistment: Report of a workshop* (pp. 37–60). The National Academies Press.

Wright, G. J. (1984). *Crosscoding military and civilian occupational classification systems* [Conference session]. Military Testing Association, Munich, Federal Republic of Germany.

Appendix

Additional Analyses, Scatterplots, and Forest Plots

Table A1

Correlations Among Mean MOS AFQT Score, Time in Service, and Job Characteristics

Measure	1	2	3	4	5	6
1. Mean AFQT score	—					
2. Mean time in service	**.49**	—				
3. Working with things	.24	**.65**	—			
4. Cognitive complexity	**.49**	**.58**	**.46**	—		
5. Unpleasant working conditions	**−.37**	−.28	−.09	**−.55**	—	
6. Fine motor control	.13	−.06	−.35	−.04	−.02	—

Note. $N = 31$. Boldface indicates $p < .05$. AFQT = Armed Forces Qualification Test; MOS = military occupation speciality.

(*Appendix continues*)

Table A2
Hierarchical Regression Analyses Showing Incremental Validity of AFQT and Time in Service for Job-Specific Performance Test Scores Performed on Correlation Matrices Corrected for Unreliability and Restriction of Range

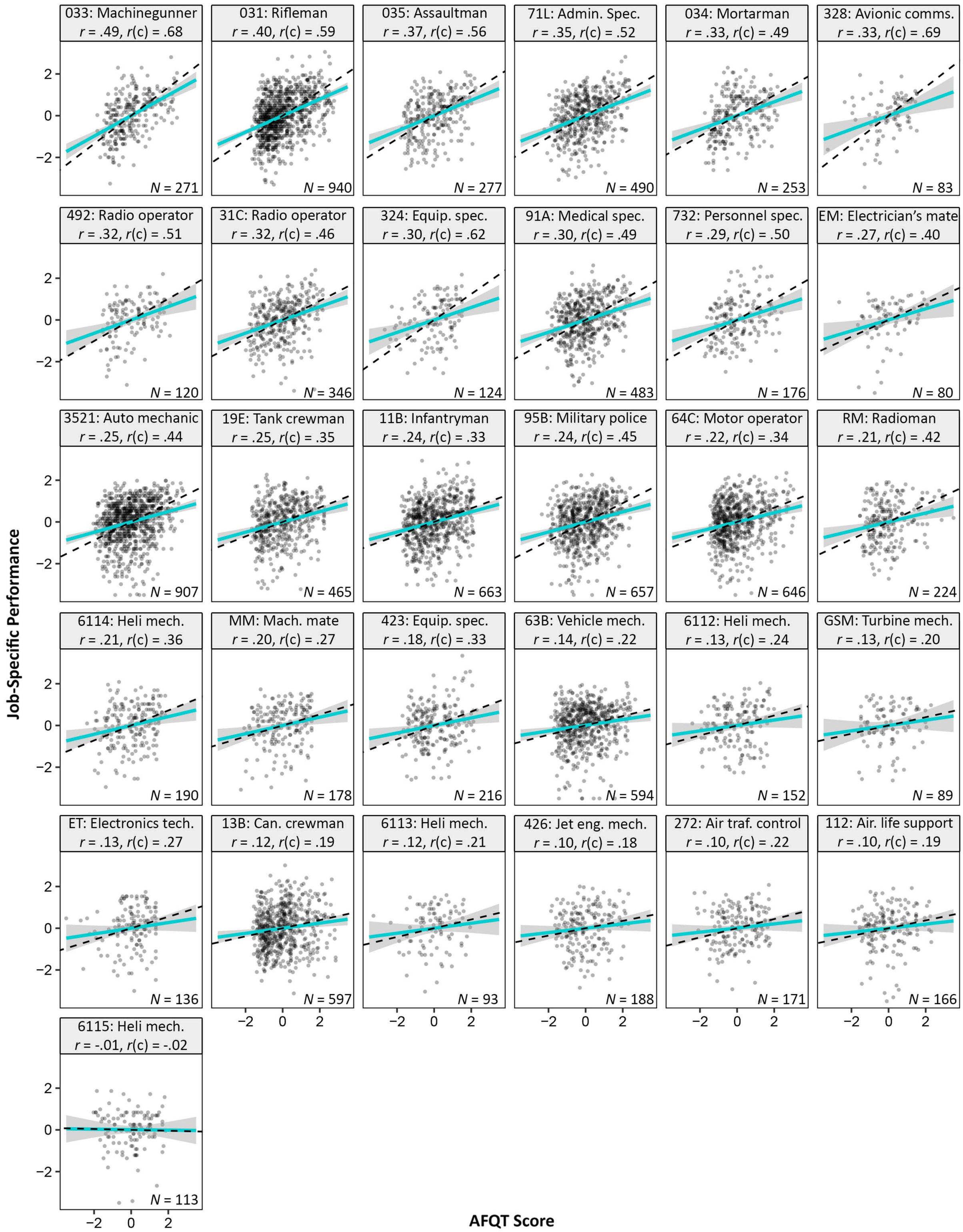
			Model 1				Model 2			
			Step 1 AFQT		Step 2 Time in service		Step 1 Time in service		Step 2 AFQT	
MOS	Description	*N* (%)	*R*	*F*	Δ*R*	Δ*F*	*R*	*F*	Δ*R*	Δ*F*
Army										
11B	Infantryman	663 (6.6)	**.33**	**79.40**	.00	0.04	.01	0.04	**.32**	**79.27**
13B	Cannon crewman	597 (5.9)	**.19**	**23.47**	**.08**	**23.36**	**.20**	**23.58**	**.07**	**23.24**
19E	Tank crewman	465 (4.6)	**.35**	**63.23**	**.05**	**21.10**	**.20**	**19.21**	**.20**	**65.21**
31C	Single-channel radio operator	346 (3.4)	**.46**	**90.59**	.00	0.21	.04	0.61	**.42**	**45.3**
63B	Light wheel vehicle mechanic	594 (5.9)	**.22**	**31.31**	.01	1.48	.02	0.17	**.21**	**16.41**
64C	Motor transport operator	646 (6.4)	**.34**	**85.45**	.00	1.22	.05	1.45	**.29**	**85.06**
71L	Administrative specialist	490 (4.9)	**.52**	**178.74**	**.01**	**6.30**	.01	0.02	**.52**	**187.05**
91A	Medical specialist	483 (4.8)	**.49**	**149.15**	.00	2.37	.01	0.03	**.48**	**151.90**
95B	Military police	657 (6.5)	**.45**	**167.75**	.01	3.76	**.08**	**4.22**	**.38**	**166.93**
Marines										
031	Rifleman	940 (9.3)	**.59**	**497.07**	**.08**	**165.08**	**.35**	**134.21**	**.32**	**538.04**
033	Machinegunner	271 (2.7)	**.68**	**225.69**	**.02**	**19.48**	**.24**	**16.78**	**.46**	**229.63**
034	Mortarman	253 (2.5)	**.49**	**79.25**	**.10**	**40.47**	**.36**	**38.46**	**.23**	**81.41**
035	Assaultman	277 (2.7)	**.56**	**124.53**	**.02**	**12.01**	**.22**	**13.67**	**.36**	**121.85**
3521	Organizational automotive mechanic	907 (9.0)	**.44**	**214.98**	**.07**	**83.93**	**.21**	**43.61**	**.30**	**262.39**
6112	Helicopter mechanic CH-46	152 (1.5)	**.24**	**9.08**	**.49**	**155.86**	**.67**	**122.23**	**.06**	**29.14**
6113	Helicopter mechanic CH-53	93 (0.9)	**.21**	**4.17**	**.56**	**125.39**	**.72**	**100.68**	**.05**	**16.94**
6114	Helicopter mechanic U/AH-1	190 (1.9)	**.36**	**27.84**	**.21**	**54.77**	**.43**	**43.68**	**.14**	**38.24**
6115	Helicopter mechanic CH-53 E	113 (1.1)	.02	0.05	**.39**	**22.77**	**.41**	**22.65**	.00	0.32
Air Force										
112	Aircrew life support specialist	166 (1.6)	**.19**	**5.93**	**.22**	**25.60**	**.41**	**32.46**	.00	0.14
272	Air traffic control operator	171 (1.7)	**.22**	**8.57**	**.17**	**20.82**	**.30**	**17.17**	**.09**	**12.09**
324	Precision measuring equipment specialist	124 (1.2)	**.62**	**76.76**	**.03**	**6.55**	**.42**	**26.35**	**.23**	**49.88**
328	Avionic communications specialist	83 (0.8)	**.69**	**72.96**	**.09**	**26.55**	**.34**	**10.71**	**.44**	**98.88**
423	Aerospace ground equipment specialist	216 (2.1)	**.33**	**26.57**	**.08**	**15.33**	**.28**	**18.29**	**.13**	**23.48**
426	Jet engine mechanic	188 (1.9)	**.18**	**6.00**	**.13**	**13.44**	**.21**	**8.69**	**.10**	**10.70**
492	Information systems radio operator	120 (1.2)	**.51**	**40.61**	**.13**	**32.00**	**.49**	**37.90**	**.15**	**34.59**
732	Personnel specialist	176 (1.7)	**.50**	**59.28**	**.17**	**61.70**	**.47**	**48.40**	**.20**	**71.19**
Navy										
EM	Electrician's mate	80 (0.8)	**.40**	**15.03**	**.14**	**14.55**	**.40**	**14.96**	**.14**	**14.62**
ET	Electronics technician	136 (1.3)	**.27**	**10.94**	**.05**	**3.95**	**.28**	**11.01**	.04	3.88
GSM	Gasoline turbine mechanic	89 (0.9)	.20	3.55	**.16**	**8.61**	**.26**	**6.19**	**.10**	**5.93**
MM	Machinist's mate	178 (1.8)	**.27**	**13.30**	**.09**	**12.43**	**.22**	**8.95**	**.14**	**16.84**
RM	Radioman	224 (2.2)	**.42**	**46.46**	**.04**	**9.78**	**.26**	**16.54**	**.40**	**38.72**

Note. MOS = military occupation specialty. AFQT = Armed Forces Qualification Test (percentile score); CH = cargo helicopter; U/AH = utility/attack helicopter. Boldface indicates $p < .05$.

(*Appendix continues*)

Figure A1

Scatter Plots Between AFQT Percentile Scores (x-Axis) and Job-Specific Performance (y-Axis)

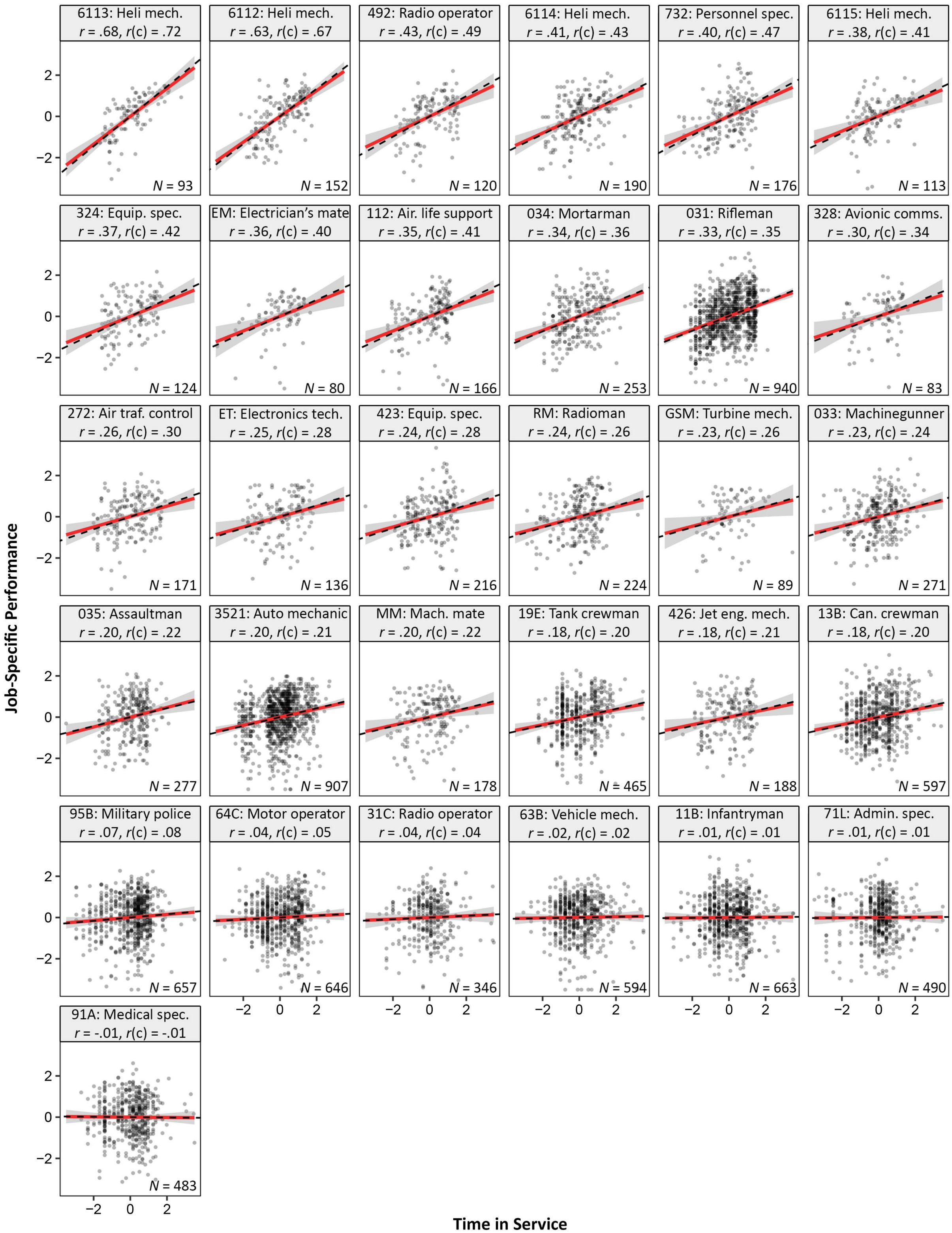


Note. Panel labels are 31 MOS codes (refer to Table 1), correlations, and corrected correlations *r*(c), ordered from strongest to weakest observed regression slope magnitude, row-wise from left to right (with corresponding 95% confidence bands). The dashed line represents the regression slope magnitude after correction for criterion unreliability and AFQT restriction of range. AFQT = Armed Forces Qualification Test; MOS = military occupation specialty. See the online article for the color version of this figure.

(*Appendix continues*)

Figure A2

Scatterplots Between Standardized Time in Service (Log Months, x-Axis) and Job-Specific Performance (y-Axis)

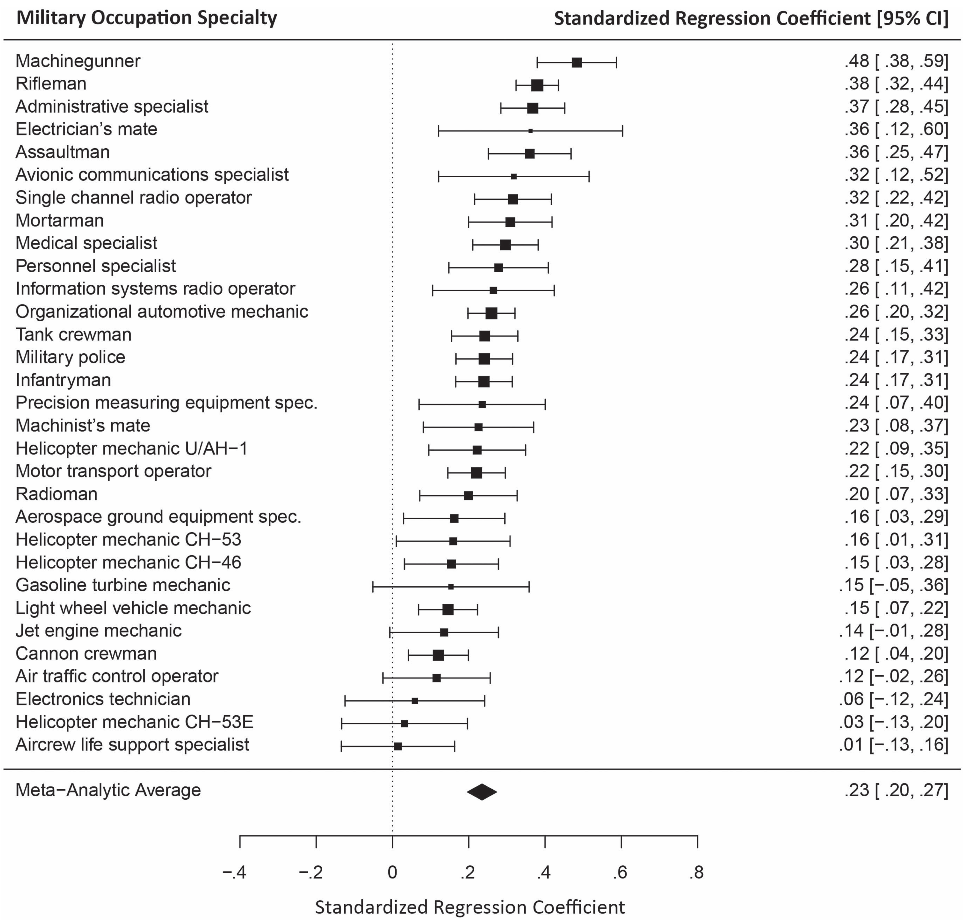


Note. Panel labels are 31 MOS codes (refer to Table 1), correlations, and corrected correlations $r(c)$, ordered from strongest to weakest observed regression slope magnitude, row-wise from left to right (with corresponding 95% confidence bands). The dashed line represents the regression slope magnitude after correction for criterion unreliability. MOS = military occupation specialty. See the online article for the color version of this figure.

(*Appendix continues*)

Figure A3

Forest Plot Depicting the Relationship Between AFQT Percentile Scores and Job-Specific Performance for 31 Military Occupation Specialties (MOSs)

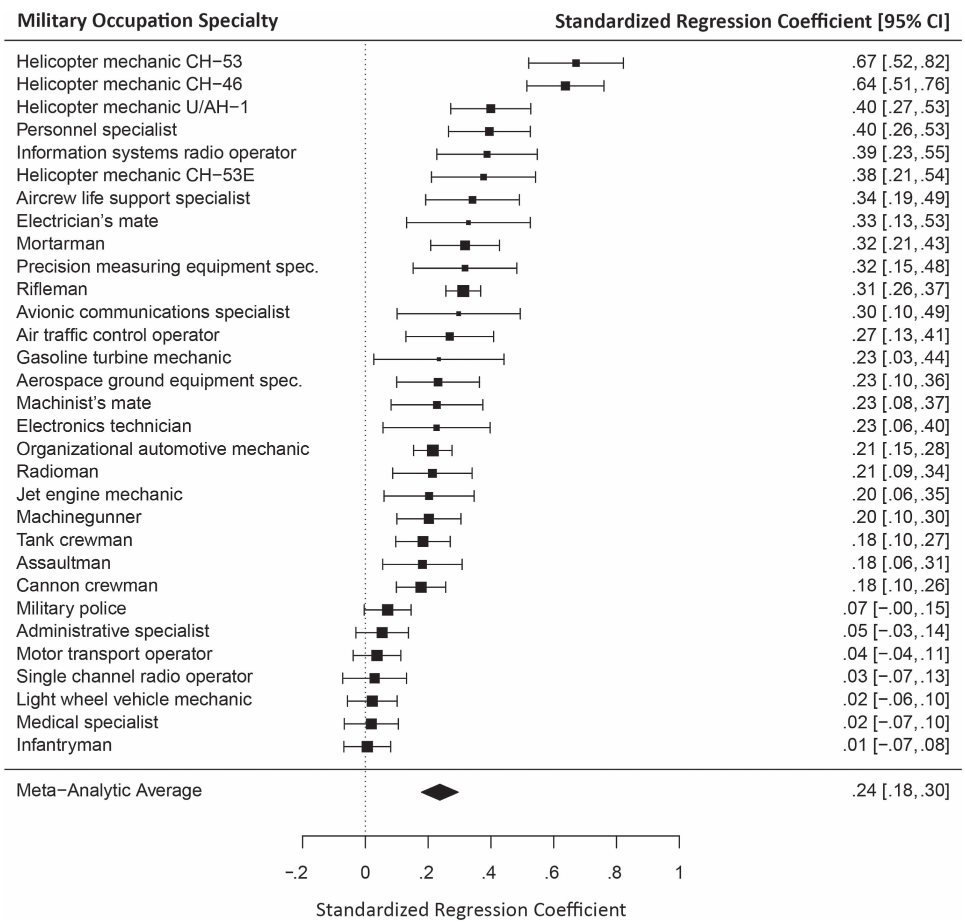


Note. Error bars represent 95% confidence intervals. Square size represents the effect's weight (inverse of sampling error variance and heterogeneity). Center of the diamond on the bottom row represents the meta-analytically weighted average effect, with the lateral tips representing its respective 95% confidence interval. AFQT = Armed Forces Qualification Test; CI = confidence interval; U/AH = utility/attack helicopter; CH = cargo helicopter.

(*Appendix continues*)

Figure A4

Forest Plot Depicting the Relationship Between Time in Service (Log Months) and Job-Specific Performance for 31 Military Occupation Specialties (MOSs)

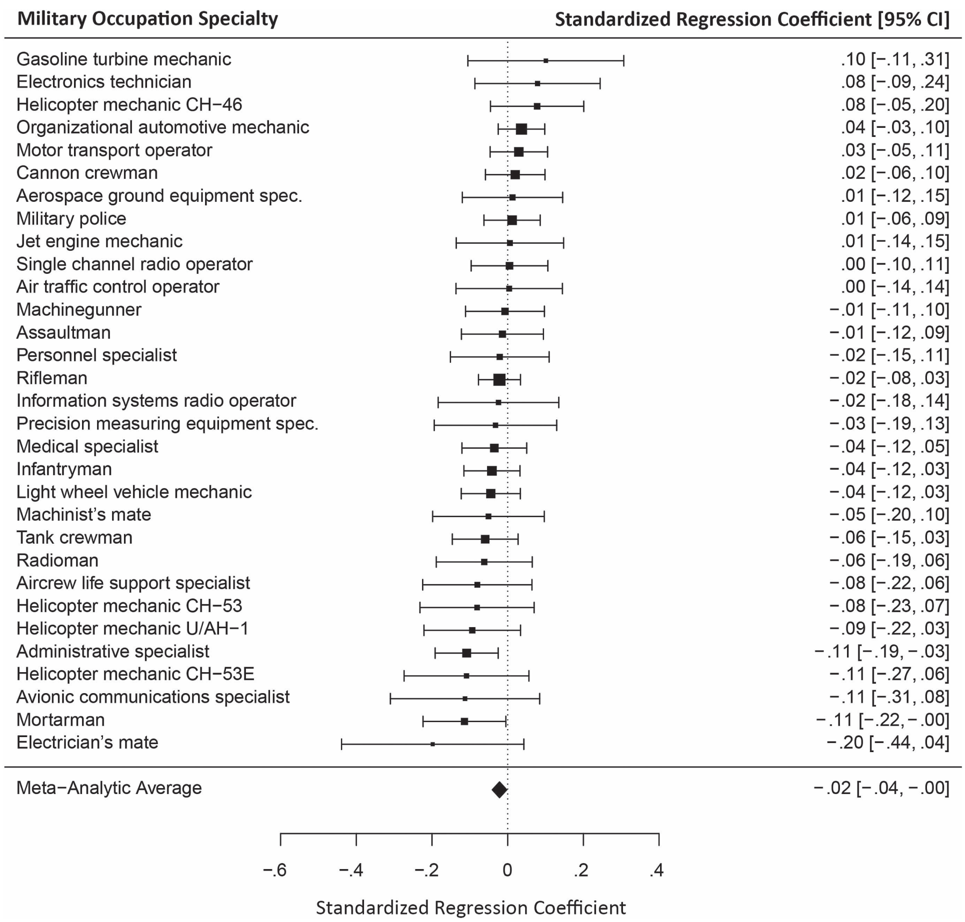


Note. Error bars represent 95% confidence intervals. Square size represents the effect's weight. The diamond on the bottom row represents the meta-analytically weighted average effect. CI = confidence interval; CH = cargo helicopter; U/AH = utility/attack helicopter.

(*Appendix continues*)

Figure A5

Forest Plot Depicting the Effect of the AFQT Percentile Score × Time in Service Interaction Term on Job-Specific Performance for Military Occupation Specialties (MOSs)



Note. Error bars represent 95% confidence intervals. Square size represents the effect's weight. The diamond on the bottom row represents the meta-analytically weighted average effect. AFQT = Armed Forces Qualification Test; CI = confidence interval; CH = cargo helicopter; U/AH = utility/attack helicopter.

Received March 4, 2022
Revision received August 8, 2023
Accepted August 9, 2023 ■