Journal of Public Economics 238 (2024) 105197

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Brains versus brawn: Ordinal rank effects in job training

Alexander J. Chesney [a,1], Scott E. Carrell [b,1,*]

[a] United States Air Force
[b] Department of Economics, University of Texas at Austin, United States of America

ARTICLE INFO

Keywords:
social interactions
Rank effects
Job training

ABSTRACT

This paper analyzes ordinal rank across cognitive and physical ability within an initial job training program. Using a rich administrative dataset and conditional random assignment of trainees to peer groups, we test whether rank effects vary across contemporaneous training and long-term career outcomes. We find cognitive ordinal rank, measured by an individual's score on the Armed Forces Qualification Test (AFQT), has a meaningful impact on completing initial training into the U.S. Air Force (USAF). This ranking also affects occupational specialization for trainees that arrive without a preassigned occupation. We also show physical ordinal rank, measured by an individual's initial fitness score, affects job training performance. Both sets of ranking effects impact behavioral misconduct outcomes and vary by gender. Finally, the interaction between cognitive and physical ordinal ranking has multiplicative effects on a limited set of outcomes.

1. Introduction

The literature on social interactions has primarily focused on using mean peer characteristics to explain individual outcomes at school and work (Sacerdote, 2014; Epple and Romano, 2011). Research by Hoxby (2000), Duflo et al. (2011), Booij et al. (2016) have also highlighted the importance of the ability distribution in human capital when analyzing non-linear and heterogeneous peer effects. This work has inspired researchers to explore alternatives to the workhorse linear-in-means specification. One popular approach leverages ordinal ranking, rather than mean ability, within peer groups to explain differences in outcomes caused by social spillovers.

In a hypothetical scenario, two individuals, A and B, of equal ability are randomly assigned to separate peer groups. Due to the naturally occurring variation in the distribution of peer human capital, individual A acquires a higher ordinal rank in his group than B. These individuals will subsequently achieve different outcomes through potential behavioral (e.g. differences in self-concept and development of non-cognitive skills) and environmental channels (e.g. differences in investment from teachers, supervisors, and family members).[2] As a result, individuals of equal ability may face substantially different outcomes based on the pure chance of possessing a higher rank in their assigned peer group.

In the described scenario, identifying the casual effect of moving up (or down) in ordinal rank exploits idiosyncratic sampling variation in cohort composition. Furthermore, rank must be exogenous to the individual. Specifically, subjects must be unable to influence placement into groups where they would enjoy higher ranking, and differences in the ability distribution across groups cannot be systematic. Lastly, in order to separate this ordinal ranking effect from other potential confounds, such as average peer ability and other environmental influences, researchers commonly include controls at the level of ranking (Denning et al., 2023).[3]

This methodology has been widely used to study the impact of ordinal rank on children and early adolescence. The most prominent studies focus on initial cognitive ranking, as measured by early test

* Corresponding author.
E-mail address: scott.carrell@austin.utexas.edu (S.E. Carrell).

[1] **Disclaimer:** The authors thank William Fischer and Brian Davis for their generous support in completing this research. A very special thanks to Tyler Danley for his critical assistance with data collection. The opinions expressed herein reflect the personal views of the authors and do not reflect the views of the United States Air Force or the Department of Defense. Research data were derived from an approved source within the Department of Defense. The study protocol was approved by the University of California-Davis Institutional Review Board in compliance with all applicable Federal regulations governing the protection of human subjects. The authors are responsible for all errors.

[2] These channels are derived from behavioral and social science research by Tversky and Kahneman (1974) on heuristics, Merton (1968) on the "Matthew Effect" and Marsh and Parker (1984), Marsh (1987) on the "big-fish–little-pond effect" (BFLPE). The most popular explanation, BFLPE theory, describes individuals obtaining higher self-perceived skill and self-concept when comparing themselves to less skilled individuals.

[3] For example, if ranking is measured at the classroom level, then a researcher should control for classroom level fixed effects since the variation in the treatment occurs within, as oppose to across, the classroom level. This naturally brings up questions on the remaining variation. As done by Elsner and Isphording (2018), we provide a detailed description of the identifying variation in the methodology section of this paper.

https://doi.org/10.1016/j.jpubeco.2024.105197
Received 20 July 2022; Received in revised form 19 April 2024; Accepted 24 July 2024
Available online 9 August 2024

scores in elementary and primary school, and the subsequent effects on future academic and job performance (Murphy and Weinhardt, 2020; Denning et al., 2023). Moreover, Kiessling and Norris (2022) demonstrate how an increase of student ordinal ranking also improves mental health outcomes. Other studies in the literature have shown the impact of rank formation is not limited to early childhood. Researchers have found ordinal ranking within high school impacts the likelihood of engaging in risky and socially deviant behavior (Elsner and Isphording, 2018) and the probability of pursuing postsecondary education (Elsner and Isphording, 2017). Furthermore, Elsner et al. (2021) demonstrate that rank formation amongst peer groups in college continue to impact academic performance and degree specialization.

However, there is a lacuna in the literature on whether ordinal ranking outside of traditional academic settings and later in life contributes to individual outcomes. In this paper, we analyze peer formation of young adults entering a job training program in preparation for a new career in the United States Air Force (USAF). We provide empirical evidence showing ordinal rank continues to have a meaningful impact across a wide menu of training and career outcomes. This is important since 52.4 percent of 2021 U.S. high school graduates who do not attend college immediately enter the workforce (U.S. Bureau of Labor Statistics, 2022a), while social interactions in the workplace impact workers' lives on many dimensions as described by Mas and Moretti (2009), Dahl et al. (2014), Cornelissen et al. (2017). With respect to ordinal rank effects at work, Brown et al. (2008), Card et al. (2012) have analyzed the relationship between salary ranking on job satisfaction and overall well-being, however, to our knowledge, there are no previous studies examining ordinal rank in job training programs. This is a significant void given recent research by Opper et al. (2022) who find an individual trained in a cohort with a mean labor market history that is one standard deviation above the average is 15 p.p. more likely to be employed afterwards relative to groups one standard deviation below.

We analyze ordinal rank effects by leveraging a unique research setting at Basic Military Training (BMT) for the USAF. This introductory program is a formative time for new enlisted trainees. BMT acquaints young adults with the structure and culture of the military, while preparing individuals for follow-on job training in their new careers across a wide spectrum of occupations as described in Fig. 1. This two month intensive training program offers an ideal setting to study rank effects for several reasons. First, new trainees are exogenously assigned to peer groups, known as flights, of 50 people on average. Trainees are in close proximity of one another throughout the entire program. Conditional on key observables, we also show assignment to flight is almost-as-good as random, which has been used in only a limited number of ordinal rank studies (Elsner et al., 2021; Bertoni and Nisticò, 2023). This feature allows us to rule out the possibility of trainees self-selecting into groups where they would enjoy higher rank.

Second, our research setting provides a wide menu of short and long-term outcomes for over 200,000 enlisted trainees entering military service from 2011–2017. For the short-term outcomes, we examine whether initial cognitive ranking within flight, measured by prior test scores on the Armed Forces Qualifying Test (AFQT), impacts completion of BMT. Conditional on own ability, we find that moving from being ranked last to being ranked in the 25th percentile increases the probability of BMT completion by roughly 1.3 percentage points. With a historic attrition rate of 6 percent, though sizeable, this effect is about one-third of that that found by Denning et al. (2023) for third grade retention. In addition to rank affecting the likelihood of graduation, we also find that moving from being ranked in the 75th percentile to the top of the group increases the likelihood of becoming an "honor" graduate by over approximately 3.3 percentage points. This effect is quite large given a 9.1 percent mean honor graduate rate. This finding is important since we demonstrate how honor graduate status has a meaningful correlation with reducing downstream costs for the organization, including trainee academic performance in follow-on technical training.

Third, the majority of research in the rank literature hinges on the rack-and-stack of individuals based on *cognitive* aptitude scores in reading, arithmetic, and STEM related subjects, yet proxying human capital with academic test scores alone appears incomplete by some researchers' standards (Heckman and Zhou, 2022) as social (Deming, 2017) and non-cognitive skills (Edin et al., 2022) have become ever more important in the modern labor force. There lies an unanswered question of whether ordinal ranking on measures other than academic ability have the same, or greater, impact on contemporaneous and long-term outcomes. For example, Pagani et al. (2021) finds ordinal ranking of the Big Five personality traits, specifically conscientiousness, has a positive and sizeable impact on early academic achievement. Additionally, work by Chanal et al. (2005) suggests the importance of fitness ranking effects when analyzing physical self-concept among gymnastic students, while Kim (2021) finds a strong meaningful correlation between body mass index (BMI) ordinal ranking and measures of life-satisfaction.

Our paper builds on this framework by demonstrating how physical ordinal rank affects individual outcomes. Specifically, we find ordinal ranking based on fitness test scores has little impact on BMT graduation, save for the lowest ranked individuals. However, evidence suggests physical fitness rank has a meaningful impact on graduating with honors. For example, moving from either being ranked last to being ranked in the 25th percentile or from the 75th percentile rank to the top of the group increases the likelihood of graduating with honors by over two percentage points. Additionally, we analyze how these effects differ between men and women, who are segregated into gender specific peer groups. For example, we find almost no impact of physical ranking on the likelihood of men completing BMT; however, physical rank is positive and significantly related to BMT completion for women. Moreover, we also find that cognitive and physical rank are negatively correlated for men and positively correlated for women.[4]

Fourth, we analyze the long-run behavioral misconduct and disciplinary outcomes of BMT graduates in follow-on technical training and during an enlistee's first 36 months in their career. Previous research has found a strong link between disruptive behavior and class rank. Cicala et al. (2017) finds a 50 percentile decrease in rank among 5th to 6th grade schoolmates is associated with a nearly 2.5 p.p. increase in the probability of a serious behavioral incident. Similarly, Elsner and Isphording (2018) shows a one decile increase in high school ordinal rank leads to a 1.4, 1.2, and 0.6 p.p. decrease in the probability of underage drinking, smoking, and unprotected sexual activity. We find effects similar in magnitude on whether an individual receives disciplinary action for behavioral misconduct. This is important because individuals who exit military service with a less-than-honorable (LTH) discharge for minor misconduct risk losing access to critical veteran's benefits (e.g. healthcare, housing, and education) and potentially hindering their civilian labor market opportunities (McClean, 2021).

Finally, our setting provides a novel opportunity to examine the mechanisms for how ordinal ranking improves outcomes. With respect to environmental channels, we show that individuals with higher physical ordinal rankings are more likely to be chosen for early leadership roles during BMT, garnering extra responsibilities and greater focus from BMT personnel. On the other hand, the impact of cognitive ordinal rankings shows a gender disparity, with minimal effects on men but a significant influence on women's leadership opportunities during training. This distinction underscores the complex interplay of physical and cognitive rankings in shaping environmental and attention dynamics.

[4] Elsner et al. (2021), Murphy and Weinhardt (2020) also demonstrate rank effects vary across gender. Differential responses to treatments, such as academic probation, has also been shown in previous studies as by Lindo et al. (2010).

AFSC	Description	AFSC	Description	AFSC	Description	AFSC	Description
1A0	In-Flight Refueling Spc	1T0	SERE Specialist	2T2	Air Trans	3N1	Regional Band
1A1	Flight Eng	1U0	Sensor Ops	2T3	Vehicle Mgmt	3N2	Premier Band – The USAF Band
1A2	Aircraft Ldm	1U1	RPA pilot	2W0	Munitions Sys	3N3	USAF Academy Band
1A3	Airborne Mission Sys Spc	1W0	Weather	2W1	Aircraft Armament Sys	3P0	Security Forces
1A6	Flight Attendant	1Z1	Pararescue	2W2	Nuclear Weapons	4A0	Health Services Mgmt
1A8	Airborne ISR	1Z2	Combat Control	3D0	Cyberspace Ops	4A1	Medical Materiel
1A9	Spc Mission Aviator	1Z3	TACP	3D1	Client Sys	4A2	Biomedical Equip
1B0	Cyber Warfare Operations Supt	1Z4	Special Recon	3E0	Facility Sys	4B0	Bioenvironmental Eng
1B4	CW Ops	2A0	Avionics	3E1	HVAC, Refrigeration	4C0	Mental Health Svc
1C0	Aviation Rsc Mgmt	2A2	SOF/PR Integrated Comm/Nav/Mission Sys	3E2	Heavy Repair	4D0	Diet Therapy
1C1	Air Traffic Control	2A3	Fighter/RPA Maint	3E3	Structural	4E0	Public Health
1C3	C2 Ops	2A5	Airlift/Special Mission Aircraft Maint	3E4	Infrastructure Sys	4H0	Cardiopulmonary Lab
1C5	C2 Battle Mgmt Ops	2A6	Aircraft Sys	3E5	Engineering	4J0	Physical Medicine
1C6	Space Sys Ops	2A7	Aircraft Fabrication	3E6	Ops Mgmt	4M0	Aerospace and Ops Physiology
1C7	Airfield Mgmt	2A8	Mobility AF Integrated Comms/Nav/Mission Sys	3E7	Fire Protection	4N0	Aerospace Medical Svc
1C8	Radar, Airfield, and Weather Sys	2A9	Bomber/Spc Integrated Comms/Nav/Mission Sys	3E8	Explosive Ordnance Disposal	4N1	Surgical Svc
1N0	Intelligence	2F0	Fuels	3E9	Emergency Mgmt	4P0	Pharmacy
1N1	Geospatial Intelligence	2G0	Logistics Plans	3F0	Personnel	4R0	Diagnostic Imaging
1N2	SIGINT	2M0	Missile and Space Sys Maint	3F1	Services	4T0	Medical Lab
1N3	Cryptologic Language Analyst	2P0	Precision Measurement Equipment Lab	3F2	Education and Training	4V0	Ophthalmic
1N4	Fusion Analyst	2R0	Maint Mgmt	3F3	Manpower	4Y0	Dental
1N7	HUMINT Spc	2R1	Maint Mgmt Prod	3F4	Equal Opportunity	5J0	Paralegal
1N8	Targeting Analyst	2S0	Materiel Mgmt	3F5	Administration	5R0	Religious Affairs
1P0	Aircrew Flight Equip	2T0	Traffic Mgmt	3H0	Historian	6C0	Contracting
1S0	Safety	2T1	Ground Trans	3N0	Public Affairs	6F0	Financial Mgmt and Comptroller

Fig. 1. Sample list of Air Force Specialty Codes (AFSC).
Data derived from https://www.airforcemag.com/article/2021-usaf-ussf-almanac-specialty-codes/ Current as of September 30, 2020.

As an alternative mechanism, we delve into how ordinal rankings influence self-perception and career occupational choices among BMT trainees. Building on previous research by Elsner et al. (2021) that demonstrated how ordinal ranking within teaching sections altered college students' decisions to specialize and take additional classes within a major, we find similar dynamics at play in our research setting; thus, complementing the rich literature on occupation choice (Antonovics and Golan, 2012; Rothstein and Rouse, 2011). We show trainees that arrive to BMT without a preassigned job are more likely to be matched with occupations of a higher average ability based on ordinal ranking. Specifically, we find that a 1-standard deviation increase in cognitive ordinal ranking leads to a 6% of a standard deviation increase in the average AFQT of the occupational assignment.[5] For physical ranking, we estimate small and insignificant effects. This finding contrasts with previous research by Jones and Kofoed (2020) who find peers have little impact on occupation choice. Instead, our findings suggest cognitive ordinal ranking enhances self-perception and motivates individuals to select into more challenging occupations.

The policy implications from the ordinal rank literature are somewhat unclear. While Bhattacharya (2009) shows that heterogeneous peer effects across groups could lead to a pareto improvement in outcomes through the systematic reassignment of individuals, efforts to improve outcomes through intentional peer formation have proven to be difficult due to unforeseen factors (Carrell et al., 2013). Our results suggest this problem may be even more complex than previously known, due to interaction effects across multiple rankings. In our setting, we find physical and cognitive ranking have an overall negative interaction effect for program completion. On the contrary, these rank interactions act in unison for determining honor graduate status, leading to potential superstar effects (Rosen, 1981).[6] As such, policy makers wishing to use ordinal rank to form a deliberate assignment process to improve outcomes need to take into account the multidimensional complexity of this optimization problem. That is, pulling on one lever may inadvertently push on another, resulting in unintended negative consequences.

The remainder of this paper is structured as follows: Section 2 explains the background of BMT; Section 3 describes the data; Section 4 provides the methodology and identification; Section 5 provides results; Section 6 discusses; and Section 7 concludes.

[5] After BMT training, all graduates attend additional training for their assigned job in the Air Force. Jobs are classified into what are referred to as Air Force Specialty Codes (AFSC's). For this analysis we examine the rank effects on the average AFQT score for all individuals who are assigned to each AFSC.

[6] These results are supported by Weingarten et al. (2018) who find people are sensitive to success (and failure) across multiple reference points, such as fitness and academic goals.

2. Background

In order to enter military service within the USAF, all enlistees must complete a nearly two month intensive training program known as BMT at Lackland AFB in San Antonio Texas. Prior to attending BMT, the application process to join the USAF starts at a local recruiting office. Potential enlistees undergo a thorough background check that verifies citizenship and criminal/financial history. Afterwards, they receive a health physical exam at the nearest Military Entrance and Processing Stations (MEPS). They are also required to take a standardized cognitive test known as the Armed Services Vocational Aptitude Battery (ASVAB). Scores on the exam will be used in determining eligibility for military service and occupational assignments (National Academies of Sciences, Engineering, and Medicine, 2021).

Upon passing this rigorous screening process and signing a contract to enlist in either a specific occupation or generalized specialty[7], a trainee will be assigned to a BMT weekly cohort composed of 700 people on average. In our study, we analyze the universe of USAF trainees from 2011–2017 in which approximately 34,000 trainees attended annually. As illustrated in Fig. 10, trainees arrive all across the world and will typically report to their assigned MEPS station on the Monday prior to BMT starting for final screening and paperwork. The following day, trainees are flown to the San Antonio Airport and subsequently transported via bus to BMT at Lackland AFB. Upon arrival, trainees are then placed into same gender segregated peer groups, known as flights, on a first-come, first-served basis.

BMT is designed to prepare enlisted trainees for follow-on technical training in their assigned occupation and instill critical disciplinary habits required for a successful career in the military. A summary of the two month program's syllabus is provided in Fig. 2. Throughout the program, trainees are led by their assigned instructors, who are responsible for daily training and instilling important career principles.[8] Each trainee is responsible for learning important organizational coursework, passing inspections and building physical stamina. At the end of BMT, trainees take a series of physical and cognitive exams required for graduation.

Upon completion of BMT, graduates are then sent to follow-on training assignments. The location and duration of technical training depend on the assigned occupation, known as Air Force Specialty

[7] See Cullen et al. (2024) for further discussion about optimal policies for occupation assignment in BMT.

[8] Trainees also receive instruction from other members of BMT staff in coursework such as Sexual Assault Prevention and Response, Combat Arms Training and Maintenance, and many other modules as listed in Fig. 2.

Fig. 2. Description of USAF BMT training program in 2015. NOTE: TYPICALLY DUTY HOURS WILL BE FROM 0545 (LIGHTS ON) TO 2100 (LIGHTS OUT)

Code (AFSC), which is described in Fig. 1.[9] Many of these positions translate to civilian jobs. Using data from the Defense Manpower Data Center and classifications derived from the Bureau of Labor Statistics (BLS) Occupational Outlook Handbook, Fig. 3 provides a breakdown of careers trainees serve under. Despite vast heterogeneity in these occupations, BMT provides a common experience that all USAF enlisted trainees share before they arrive at their first job site at a military installation.

There has been considerable research into what makes a successful military trainee (Marrone, 2020). Predictive information includes ability tests, medical evaluations, background interviews, prior education, and criminal records. There are recent efforts to incorporate non-cognitive personality tests (Trent et al., 2020) and web-based vocational interest tools (Johnson et al., 2020) to improve occupation selection and retention. Chesney et al. (2024) also explains that the

[9] For example, an enlistee assigned to a Security Forces (3PO) AFSC will undergo follow-on training for several months learning weaponry, laws, installation security, and other police related skills. Alternatively, individuals selected to work in the weather career field (1W0) spend many months learning how to operate meteorological equipment and employ computer workstations to interrogate current and forecast atmospheric and space weather conditions (Air Force Personnel Center, 2021).

	Number	Percentage
Occupational Group		
Administrative	14,145	5.27%
Combat Specialty	720	0.27%
Construction	5,134	1.91%
Electronic and Electrical Equipment Repair	32,259	12.03%
Engineering, Science, and Technical	53,672	20.01%
Healthcare	15,409	5.75%
Human Resource Development	8,621	3.21%
Machine Operator and Production	6,750	2.52%
Media and Public Affairs	5,874	2.19%
Protective Service	36,009	13.43%
Support Service	5,603	2.09%
Transportation and Material Handling	29,623	11.05%
Vehicle and Machinery Mechanic	47,487	17.71%
Non-occupation or unspecified coded personnel	6,859	2.56%
Total enlisted personnel as of 2021	268,165	

Fig. 3. Description of USAF enlisted occupations.
Data derived from U.S. Department of Defense, Defense Manpower Data Center, March 2021 and Bureau of Labor Statistics, Occupational Outlook Handbook, Military Careers https://www.bls.gov/ooh/military/military-careers.htm.

assigned instructors, known as Military Training Instructors (MTIs), play a critical role in determining success.[10] However, the role of peer or ordinal ranking effects during this initial training program has been understudied.

This analysis naturally brings up questions on external validity. First, Fig. 3 demonstrates that trainee occupations can be compared to jobs in the civilian sector in which U.S. Bureau of Labor Statistics (2022b) provides a detailed technical report on this matter. Second, military service is an important part of the modern U.S. economy. Nearly 1.3 M individuals are actively serving in the military in which there has been a strong transition in the services demographics to become more representative of the overall population (Council on Foreign Relations, 2020). More importantly, 7% of working age adults in the U.S. are veterans (Schaeffer, 2021). Despite earlier studies by Angrist (1990) that show a substantial economic penalty for draftees, Angrist (1998) finds volunteer military service is associated with higher employment rates for veterans. Recent analysis by Greenberg et al. (2022) has also highlighted the role of military service as a driver for economic mobility, specifically for underrepresented minorities. However, military members that exit service with a LTH discharge for poor performance and minor misconduct may suffer substantially lower civilian employment opportunities and risk losing access to critical veterans' benefits for housing, health, and education (McClean, 2021). Finally, the average age of a new Air Force enlistee is around twenty years. As such, the population in our sample is more similar to a college campus. In terms of race/ethnicity demographics, the Air Force is relatively comparable to the US as whole, with 17 percent Hispanic, 15 percent Black, and 5 percent Asian, though, women are relatively underrepresented (SAFPA, 2021).

Table 1
USAF basic military trainees entering from October 2011–December 2017.

	Full	Men	Women
Women	0.227	0	1
	(0.419)	–	–
Age	20.62	20.58	20.76
	(3.201)	(3.130)	(3.428)
Black	0.190	0.172	0.250
	(0.392)	(0.377)	(0.433)
Hispanic	0.145	0.144	0.149
	(0.352)	(0.351)	(0.356)
Asian	0.049	0.049	0.047
	(0.215)	(0.216)	(0.212)
Married	0.101	0.096	0.116
	(0.301)	(0.295)	(0.320)
Some College	0.101	0.096	0.120
	(0.302)	(0.294)	(0.325)
Bachelor's Degree or more	0.045	0.041	0.060
	(0.208)	(0.198)	(0.237)
No Health Issues	0.741	0.749	0.713
	(0.438)	(0.434)	(0.452)
Waiver Required	0.094	0.094	0.094
	(0.292)	(0.292)	(0.291)
Active Duty	0.794	0.813	0.730
	(0.405)	(0.390)	(0.444)
Guaranteed Job (Active Only)	0.656	0.663	0.631
	(0.475)	(0.473)	(0.482)
Armed Forces Qualification Test (AFQT)	69.23	70.40	65.25
	(16.01)	(15.95)	(15.57)
Initial Physical Fitness Score	65.30	67.43	58.05
	(26.66)	(26.01)	(27.54)
Observations	215,132	166,362	48,770

3. Data

Our population is the universe of USAF enlistees entering basic training from October 2011 through December 2017. We observe follow-on outcomes through 2020. Full summary statistics are described in Table 1. Data for this study were gathered from Air Education and Training Command (AETC), Air Force Recruiting Services (AFRS), and the Air Force Personnel Center (AFPC). This compiled administrative dataset includes a rich set of covariates measured prior to arrival at BMT, such as race, gender, education history, martial status, assigned occupation, medical history, and whether the trainee required a waiver to join the military. We possess exhaustive information on an individual's experience at BMT. This includes records on a trainee's assigned flight and peers, with a wide range of extensive and intensive measures of success throughout the program.

We measure cognitive ranking within BMT using predetermined scores on the AFQT. Scores on the AFQT are derived from four selected

[10] The instructors responsible for training new enlistees must also meet strict qualifications, including passing a mental health evaluation and receiving a recommendation from a superior (Bloem, 2015).

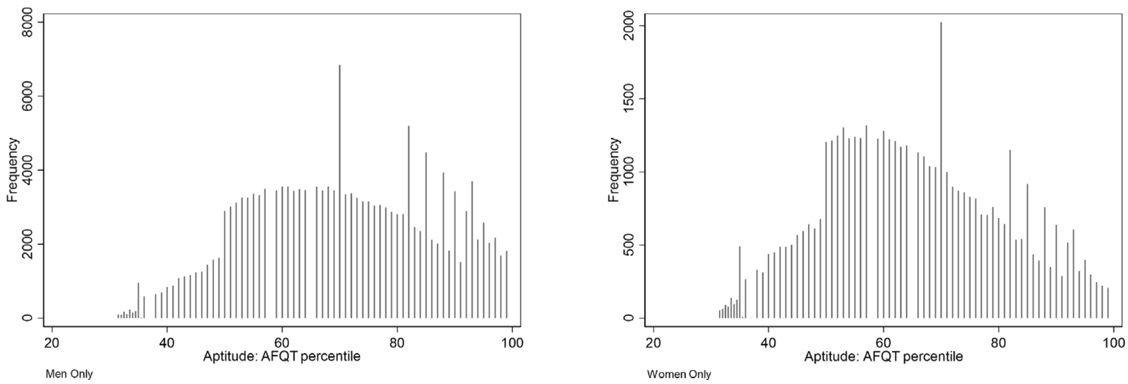

Fig. 4. AFQT Distribution.
Fig. 4 provides a distribution of AFQT scores of all incoming trainees to BMT. Fig. 5 plots ordinal cognitive rank on the *y*-axis, and demeaned AFQT score on the *x*-axis. Both graphs are separated by gender.

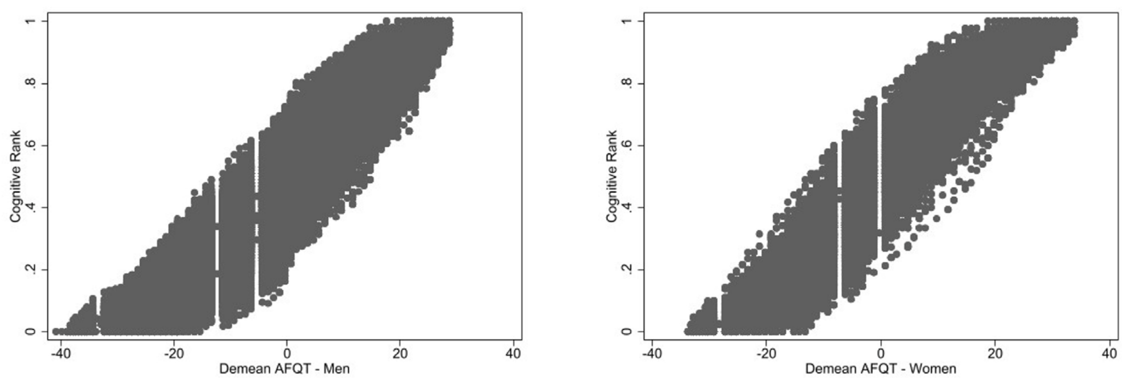

Fig. 5. AFQT variation in rank.

portions of the ASVAB in which a score is calculated and converted into a percentile-rank of 1–99. The AFQT has been used in many studies to measure cognitive ability of workers (Lang and Manove, 2011; Neal and Johnson, 1996). In our dataset, the distribution of AFQT scores is provided in Fig. 4.[11] Additionally, Fig. 5 describes how a trainee's cognitive rank within a flight can vary based on their demeaned AFQT score. For example, an individual with an average AFQT score can have a ranking that ranges from as low as 0.24 to as high as 0.72 by pure chance.

We measure physical fitness ranking using initial test results during the first week of BMT. This physical test score is scored from 0–100 based on aerobic and anaerobic measures shown in Figs. 6 and 7.[12] There are a few features of this measure that make it different than the AFQT. First, as illustrated in Fig. 8, the distribution of scores is bimodal. This is due to the scoring system having significant thresholds that lead to substantial drop offs in score if a trainee fails to meet a minimum number in a category. As a result, we elect to use dummy variables for ventiles of ability in our main specification to control for this non-linearity, as opposed to a polynomial used in other ranking studies. Second, scores are measured at the individual flight level, unlike other test scores that are collected prior to arrival. This could potentially introduce measurement error. Lastly, some individuals do not have initial fitness scores due to being on a waiver upon entry. As a result, we limit our analysis to individuals that have a valid score and only calculate the rank amongst individuals who have initial scores. We only include individuals that have valid measurements for all three categories.[13] Although imperfect, initial physical scores provide the best available proxy for fitness ability of new trainees. Referencing Fig. 9, an individual with an average physical test score can have a ranking near the bottom to as high as 0.70.

Following BMT, we have detailed training records for all active-duty trainees. We record individual's performance in follow-on technical training, specifically observing whether they experience any exam or academic block failures. Additionally, we observe whether individuals receive counseling for disciplinary issues during technical training that can arise from either poor performance or misconduct.[14]

[11] The distribution of AFQT scores for Air Force Trainee's possesses a normal distribution with two cutoffs: 31 (the lowest typical score without a waiver that a high school graduate can enter with) and 50 (the lowest GED holder score). See Greenberg et al. (2022) for a more robust discussion on the AFQT distribution and these cut-offs.

[12] The Fitness test consists of pushups, situps, run time, and waist measurement.

[13] In Appendix A.2, we provide a robustness check on these missing scores by replacing them with various values. This bounding exercise demonstrates that our originally estimated values are reasonable. Additionally, we observe ranking of other physical characteristics such as BMI, waist, and height measurements. We leverage these other ranking measures in an additional robustness check in Appendix A.3 and A.4 to show how our main specification holds.

[14] In our sample, 79% of incoming trainees are active-duty members. The remaining individuals consist of Air Force Reservists and Air National Guardsmen. For this latter group, we have incomplete data for our follow-on outcomes. As a result, we use the full sample of trainees when estimating effects on short-term outcomes and only use active-duty members when analyzing longer-term outcomes. In Appendix A.1, we analyze all outcomes using only active-duty members. Results for short-term outcomes do not differ significantly from our preferred specification with the full sample.

FITNESS ASSESSMENT CHART – MALE: AGE: < 30

Cardiorespiratory Endurance			Body Composition			Muscle Fitness			
Run Time (mins:secs)	Health Risk Category	Points	AC (inches)	Health Risk Category	Points	Push-ups (reps/min)	Points	Sit-ups (reps/min)	Points
≤ 9:12	Low-Risk	60.0	≤ 32.5	Low-Risk	20.0	≥ 67	10.0	≥ 58	10.0
9:13 - 9:34	Low-Risk	59.7	33.0	Low-Risk	20.0	62	9.5	55	9.5
9:35 - 9:45	Low-Risk	59.3	33.5	Low-Risk	20.0	61	9.4	54	9.4
9:46 - 9:58	Low-Risk	58.9	34.0	Low-Risk	20.0	60	9.3	53	9.2
9:59 - 10:10	Low-Risk	58.5	34.5	Low-Risk	20.0	59	9.2	52	9.0
10:11 - 10:23	Low-Risk	57.9	35.0	Low-Risk	20.0	58	9.1	51	8.8
10:24 - 10:37	Low-Risk	57.3	35.5	Moderate Risk	17.6	57	9.0	50	8.7
10:38 - 10:51	Low-Risk	56.6	36.0	Moderate Risk	17.0	56	8.9	49	8.5
10:52 - 11:06	Low-Risk	55.7	36.5	Moderate Risk	16.4	55	8.8	48	8.3
11:07 - 11:22	Low-Risk	54.8	37.0	Moderate Risk	15.8	54	8.8	47	8.0
11:23 - 11:38	Low-Risk	53.7	37.5	Moderate Risk	15.1	53	8.7	46	7.5
11:39 - 11:56	Low-Risk	52.4	38.0	Moderate Risk	14.4	52	8.6	45	7.0
11:57 - 12:14	Low-Risk	50.9	38.5	Moderate Risk	13.5	51	8.5	44	6.5
12:15 - 12:33	Low-Risk	49.2	39.0 *	Moderate Risk	12.6	50	8.4	43	6.3
12:34 - 12:53	Moderate Risk	47.2	39.5	High Risk	11.7	49	8.3	42 *	6.0
12:54 - 13:14	Moderate Risk	44.9	40.0	High Risk	10.6	48	8.1	41	5.5
13:15 - 13:36 *	Moderate Risk	42.3	40.5	High Risk	9.4	47	8.0	40	5.0
13:37 - 14:00	High Risk	39.3	41.0	High Risk	8.2	46	7.8	39	4.5
14:01 - 14:25	High Risk	35.8	41.5	High Risk	6.8	45	7.7	38	4.0
14:26 - 14:52	High Risk	31.7	42.0	High Risk	5.3	44	7.5	37	3.5
14:53 - 15:20	High Risk	27.1	42.5	High Risk	3.7	43	7.3	36	3.3
15:21 - 15:50	High Risk	21.7	43.0	High Risk	1.9	42	7.2	35	3.0
15:51 - 16:22	High Risk	15.5	≥ 43.5	High Risk	0.0	41	7.0	34	2.5
16:23 - 16:57	High Risk	8.3				40	6.8	33	2.0
≥ 16:58	High Risk	0.0				39	6.5	32	1.5
						38	6.3	31	1.3
						37	6.0	30	1.0
						36	5.8	≤ 29	0.0
						35	5.5		
						34	5.3		
						33 *	5.0		
						32	4.8		
						31	4.5		
						30	4.3		
NOTES:						29	4.0		
Health Risk Category = low, moderate or high risk for current and future						28	3.8		
cardiovascular disease, diabetes, certain cancers, and other health problems						27	3.5		
						26	3.0		
Passing Requirements - member *must*: 1) meet minimum value in each of						25	2.8		
the four components, *and* 2) achieve a composite point total ≥ 75 points						24	2.5		
						23	2.3		
* Minimum Component Values						22	2.0		
Run time ≤ 13:36 mins:secs / Abd Circ ≤ 39.0 inches						21	1.8		
Push-ups ≥ 33 repetitions/one minute / Sit-ups ≥ 42 repetitions/one minute						20	1.7		
						19	1.5		
Composite Score Categories:						18	1.0		
Excellent ≥ 90.0 pts / Satisfactory = 75.0 - 89.9 / Unsatisfactory < 75.0						≤ 17	0.0		

Fig. 6. Male fitness scores.
The following score guide is derived from Air Force Instruction (AFI) 36-2905, Fitness Program. It provides a detailed breakdown of how a fitness score is calculated from pushups, situps, run time, and waist measurement.

Post technical training, we follow an individual during their first 36 months in their career. We observe whether an individual receives disciplinary action at work below the threshold of a general courts martial. Specifically, we identify a disciplinary infraction if the individual receives non-judicial punishment from their commander, given an unfavorable information file, or placed on a control roster. Receiving such disciplinary action can be the result of various activities including minor misconduct at work, and any failure to maintain standards of conduct, military bearing, and integrity, both on and off duty (Air Force Instruction 36-2907, 2014).

4. Methodology

We follow a well established methodology presented by Denning et al. (2023), Murphy and Weinhardt (2020), Elsner and Isphording (2017) where ordinal rank is defined within peer group as the following percentile:

$$R = \frac{n_i - 1}{N_i - 1}$$

where n_i is the trainee's ordinal rank of ability measured by test scores in their respective flight and N_i is the total number of trainees assigned to the flight.[15] Ranking is bounded between 0 and 1. Individuals with the highest rank possess a 1, while trainees with the lowest rank are identified with a 0.

This measure naturally brings up questions of whether it is necessary, or even possible, for trainees to know their exact numerical *rank*. As expressed by Elsner and Isphording (2017), identification does not rely on knowing the precise number as long as sufficient variation in ranking exists and trainees can make assessments on their approximate position in the ability distribution. In our setting, trainees spend more time together than the typical subjects observed in primary

[15] As discussed in Denning et al. (2023) measurement error in the ability variable can be correlated with a person's rank, as such we follow the procedure outlined in Denning et al. (2023) by first taking the percentile of each person's ability in the entire population of BMT participants and then calculate the within-group percentile rank on this percentalized score.

FITNESS ASSESSMENT CHART – FEMALE: AGE: < 30

Cardiorespiratory Endurance			Body Composition			Muscle Fitness			
Run Time (mins:secs)	Health Risk Category	Points	AC (inches)	Health Risk Category	Points	Push-ups (reps/min)	Points	Sit-ups (reps/min)	Points
≤ 10:23	Low-Risk	60.0	≤ 29.0	Low Risk	20.0	≥ 47	10.0	≥ 54	10.0
10:24 - 10:51	Low-Risk	59.9	29.5	Low Risk	20.0	42	9.5	51	9.5
10:52 - 11:06	Low-Risk	59.5	30.0	Low Risk	20.0	41	9.4	50	9.4
11:07 - 11:22	Low-Risk	59.2	30.5	Low Risk	20.0	40	9.3	49	9.0
11:23 - 11:38	Low-Risk	58.9	31.0	Low Risk	20.0	39	9.2	48	8.9
11:39 - 11:56	Low-Risk	58.6	31.5	Low Risk	20.0	38	9.1	47	8.8
11:57 - 12:14	Low-Risk	58.1	32.0	Moderate Risk	17.6	37	9.0	46	8.6
12:15 - 12:33	Low-Risk	57.6	32.5	Moderate Risk	17.1	36	8.9	45	8.5
12:34 - 12:53	Low-Risk	57.0	33.0	Moderate Risk	16.5	35	8.8	44	8.0
12:54 - 13:14	Low-Risk	56.2	33.5	Moderate Risk	15.9	34	8.6	43	7.8
13:15 - 13:36	Low-Risk	55.3	34.0	Moderate Risk	15.2	33	8.5	42	7.5
13:37 - 14:00	Low-Risk	54.2	34.5	Moderate Risk	14.5	32	8.4	41	7.0
14:01 - 14:25	Low-Risk	52.8	35.0	Moderate Risk	13.7	31	8.3	40	6.8
14:26 - 14:52	Low-Risk	51.2	35.5 *	Moderate Risk	12.8	30	8.2	39	6.5
14:53 - 15:20	Moderate Risk	49.3	36.0	High Risk	11.8	29	8.1	38 *	6.0
15:21 - 15:50	Moderate Risk	46.9	36.5	High Risk	10.7	28	8.0	37	5.5
15:51 - 16:22 *	Moderate Risk	44.1	37.0	High Risk	9.6	27	7.5	36	5.3
16:23 - 16:57	High Risk	40.8	37.5	High Risk	8.3	26	7.3	35	5.0
16:58 - 17:34	High Risk	36.7	38.0	High Risk	6.9	25	7.2	34	4.5
17:35 - 18:14	High Risk	31.8	38.5	High Risk	5.4	24	7.0	33	4.3
18:15 - 18:56	High Risk	25.9	39.0	High Risk	3.8	23	6.5	32	4.0
18:57 - 19:43	High Risk	18.8	39.5	High Risk	2.0	22	6.3	31	3.5
19:44 - 20:33	High Risk	10.3	≥ 40.0	High Risk	0.0	21	6.0	30	3.0
≥ 20:34	High Risk	0.0				20	5.8	29	2.8
						19	5.5	28	2.5
						18 *	5.0	27	2.0
NOTES:						17	4.5	26	1.8
Health Risk Category = low, moderate or high risk for current and future						16	4.3	25	1.7
cardiovascular disease, diabetes, certain cancers, and other health problems						15	4.0	24	1.5
						14	3.5	23	1.0
Passing Requirements - member *must*: 1) meet minimum value in each of						13	3.0	≤ 22	0.0
the four components, ***and*** 2) achieve a composite point total ≥ 75 points						12	2.8		
						11	2.5		
* Minimum Component Values						10	2.0		
Run time ≤ 16:22 mins:secs / Abd Circ ≤ 35.5 inches						9	1.5		
Push-ups ≥ 18 repetitions/one minute / Sit-ups ≥ 38 repetitions/one minute						8	1.0		
						≤ 7	0.0		
Composite Score Categories:									
Excellent ≥ 90.0 pts / Satisfactory = 75.0 - 89.9 / Unsatisfactory < 75.0									

Fig. 7. Female fitness scores.
The following score guide is derived from Air Force Instruction (AFI) 36-2905, Fitness Program. It provides a detailed breakdown of how a fitness score is calculated from pushups, situps, run time, and waist measurement.

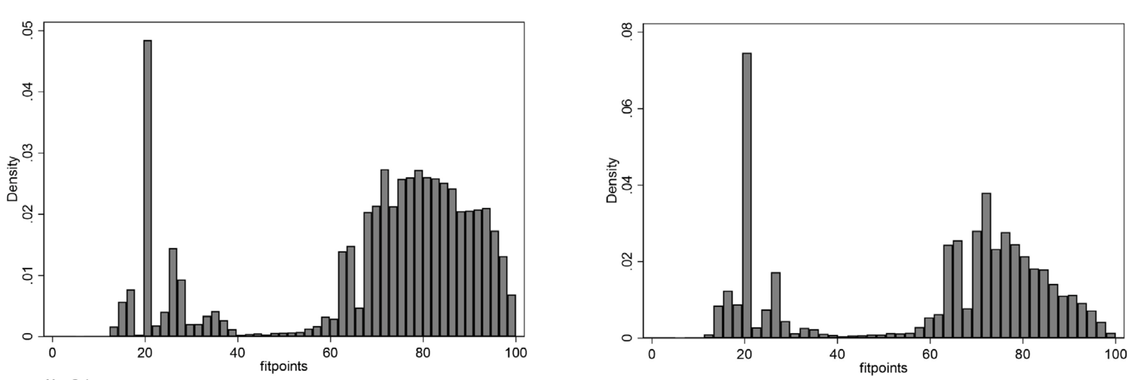

Fig. 8. Fitness score distribution.
Fig. 8 provides a distribution of physical fitness scores of all incoming trainees to BMT. Scores are derived based on performance in fitness exam using Figs. 6 and 7. Fig. 9 plots ordinal physical rank on the y-axis, and demeaned fitness score on the x-axis. Both graphs are separated by gender.

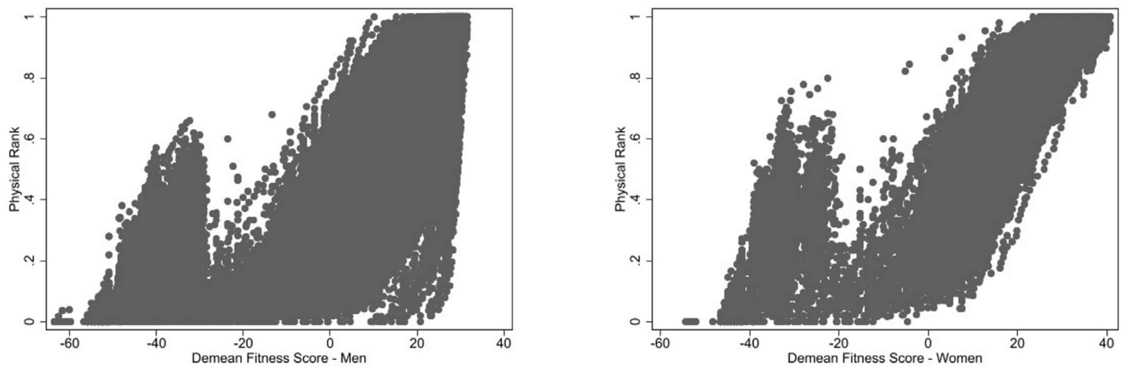

Fig. 9. Variation in fitness score rank.

school or college study groups. Additionally, trainees have intense daily interactions with their instructors and peers that allow them to make reliable judgements of where they stand in relative cognitive rank. Lastly, an initial fitness test is taken during the first week of training as a flight. Trainees can accurately perceive where they rank in some of the most salient events, such as the aerobic fitness assessment.[16]

To model trainee's outcomes based on ordinal ranking, we begin with the following education-production function described by Delaney and Devereux (2022) to explain both contemporaneous and later career outcomes:

$$Y = f(R, A, \mathbf{X}, \mathbf{F})$$

where R is the trainee's ordinal ranking in BMT, A is their measured human capital, $\mathbf{X}$ is a vector of characteristics, and $\mathbf{F}$ is a vector of flight and cohort specific attributes including peer and instructor quality. We also assume rank is additively separable from all other inputs of the production function. An implicit assumption in our identification strategy (and the ordinal rank literature in general) is that the metrics chosen to estimate ordinal rank are orthogonal to other potential metrics with which individuals compare themselves. We validate this assumption in Appendix A.3 and A.4 where we control for additional rank variable as a robustness check.[17]

The identification of the rank effect leverages the idiosyncratic variation of the ability distribution in each flight. We specifically refer to the differences in the shape across flights caused by higher moments such as the variance, skewness, kurtosis, etc. Identifying the causal effect of rank rests on two primary assumptions. First, selection into a specific flight is exogeneous to the individual. We thoroughly explore this assumption in the next section. Second, the difference in the ability distribution across flights is not driven by factors that are related to a trainee. For example, this assumption would be violated if trainees with a specific background characteristic were systematically clustered together. A straightforward way to address this concern is to control for a rich a set of predetermined characteristics. Moreover, if assignment of certain trainees to specific groups is driven explicitly by observable characteristics, such as gender or occupation, as in our research setting, including these variables as controls in our main specification alleviates this concern.

[16] Experimental work by Gill et al. (2019), Klausmann et al. (2021) that directly provides rank assignment to subjects finds similar effects as non-experimental empirical work where students must learn their rank through repeated interactions.

[17] Specifically, we additionally control for each individuals height and BMI rank and show the inclusion of these rank variables leaves our main estimates on both cognitive and physical rank largely unchanged.

4.1. Empirical model

With this framework, our preferred specification follows a similar format as Elsner and Isphording (2017) and can be written as:

$$Y_{ifc} = \alpha + \beta R_{ifc} + f(Ability_i) + X_i + \gamma_i + \phi_{fc} + \varepsilon_{ifc} \quad (1)$$

where R_{ifc} is the ordinal ranking of individual i in flight f of cohort c, $f(Ability_i)$ is a function of ventiles of ability measured by cognitive or physical test scores, X_i is a vector of attributes including race, age, education, and whether the individual required a waiver to enter BMT, γ_i is a set of randomization controls including MEPs location and assigned occupation, and ϕ_{fc} is a set of flight-by-cohort fixed effects. We cluster our standard errors at the cohort level to account for common shocks in which some cohorts may have a higher ability than others based on the time of year. In order for β to have a causal interpretation, we assume strict exogeneity of the error term ε_{ifc} with respect to the other terms, $E(\varepsilon_{ifc}|R_{ifc}, A_i, X_i, \gamma_i, \phi_{fc}) = 0$. Identification of the causal effect hinges on appropriately controlling for own ability since it is strongly related to rank, and potentially other observables. Previous studies have used higher-order polynomials to adequately capture the full extent of ability. In our setting, we use bin dummies, similar to Denning et al. (2023), since physical fitness scores possess significant non-linearity.[18]

Similar to other ordinal rank studies, we also control for fixed-effects at the level of ranking. In our setting, this is at the individual flight ϕ_{fc}. Doing so provides two important benefits. First, flight fixed-effects control for any potential confounds and common shocks at the group level that could bias our estimates. Second, the inclusion of flight fixed effects also adjusts for mean differences between peer groups. This ensures that the ability distributions of all peer groups have the same mean, making a straightforward comparison across flights. However, the ability distributions will still vary in their shape, and it is these differences in the higher moments that drive the identifying variation for ordinal ranking.

Including fixed effects at the peer group level (e.g., flight) will naturally lead a careful observer to question how much variation in rank remains. In Table 2, we provide a thorough breakdown of the standard deviation and remaining R^2 after controlling for flight fixed effects and

[18] We follow the previous literature and control for own absolute ability using ventile bin dummies. As a robustness check in columns 2 and 4 of Appendix A.3 and A.4 we re-estimated our main models while controlling for two percentage point bins of own ability. Results show the estimates largely not sensitive to this more parsimonious control, save for the honor graduate outcome, where the effect sizes decrease somewhat. We suspect this result is likely due to the small fraction of positive outcomes for this variable within each bin.

Table 2
Variation in key variables after fixed effect transformations.

	(1) Raw SD	(2) Flight FE
Cognitive Ordinal Rank	0.295	0.295
		[0.999]
Cognitive Ordinal Rank Conditional on Ability	0.084	0.051
		[0.030]
AFQT Score	16.012	15.570
		[0.946]
Fitness Ordinal Rank	0.301	0.301
		[0.999]
Fitness Ordinal Rank Conditional on Ability	0.126	0.071
		[0.055]
Fitness Score	25.417	23.755
		[0.873]
Graduation from BMT	0.240	0.237
		[0.974]
Honor Graduate from BMT	0.287	0.282
		[0.964]
No Verbal Counseling	0.481	0.474
		[0.966]
No Disciplinary Action	0.475	0.444
		[0.870]

Note: Results from the above table summarize standard deviations of predicted variables after linear regression of variables indicated on the left on sets of fixed effects. In the first column, no fixed effects are included and only the raw standard deviations are reported. In column two, entry flight fixed effects are included. Numbers in the square brackets report the share of remaining variation $(1 - R^2)$.

Table 3
Specification check - cognitive ranking on completion of BMT.

	(1) Cohort FE	(2) Cohort FE	(3) Cohort FE	(4) Cohort FE	(5) Flt FE
Cognitive Ordinal Rank	0.0259***	0.0309**	0.0309**	0.0377**	0.0317**
	(0.0076)	(0.0101)	(0.0101)	(0.0105)	(0.0103)
$\overline{AFQT_{-i}}$		0.0002	0.0002	−0.0024*	
		(0.0003)	(0.0003)	(0.0012)	
$SD(AFQT_{-i})$			−0.0002	−0.0139*	
			(0.0004)	(0.0060)	
$\overline{AFQT_{-i}}$ x $SD(AFQT_{-i})$				0.0002*	
				(0.0001)	
N	215,132	215,132	215,132	215,132	215,132
r2	0.020	0.020	0.020	0.020	0.040

Standard errors are clustered at the cohort level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) on completing BMT. Ordinal cognitive ranking is measured at the flight level based on scores on the AFQT. Each column adds an additional interaction of flight distribution characteristics. All regressions control for trainee characteristics, ventile of achievement, occupation, and MEPs station. In the first four columns, each regression controls for cohort fixed effects and gender. In the last column we control for flight fixed effects. Column 5 is our preferred specification.

prior achievement. For instance, cognitive and physical ordinal ranking have an unconditional standard deviation of 0.295 and 0.301. Given the average peer group is 50 people, a standard deviation change in ranking can equate to 14.70 and 15.05 absolute rank positions. After controlling for ability and flight fixed effects, these standard deviations are reduced to 0.051 and 0.071, respectively, which equates to a change of 2.55 and 3.55 absolute rank positions. Given how small our peer groups are compared to other research studies that typically examine rank effects in much larger class cohorts, we believe these changes in rankings are substantial.

Lastly, in Table 3 we demonstrate how our preferred specification in column 5 compares to other specifications that only control for cohort fixed effects and flight characteristics such as the mean and standard deviation of flight ability, and the interaction between mean and standard deviation. Reassuringly, estimates of the ordinal ranking effect on program completion do not change considerably across these less parsimonious specifications.

Table 4
Exogeneous assignment.

	(1) Cognitive Rank	(2) Physical Rank
Black	−0.0006	0.0002
	(0.0003)	(0.0005)
Hispanic	−0.0001	0.0009
	(0.0004)	(0.0005)
Asian	−0.0009	0.0026***
	(0.0006)	(0.0007)
Age	0.0001	0.0001
	(0.0000)	(0.0001)
Any College Plus	0.0004	0.0011*
	(0.0003)	(0.0005)
Married	0.0004	−0.0007
	(0.0004)	(0.0006)
Health Screening	−0.0003	0.0004
	(0.0003)	(0.0004)
Waiver	0.0008*	−0.0005
	(0.0004)	(0.0006)
Enlistment Contract	−0.0006	−0.0007
	(0.0003)	(0.0004)
N	215,132	211,930
r2	0.970	0.945

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results are from Eq. (2). The first column regresses cognitive ranking on individual characteristics. The second column regresses physical ranking on individual characteristics. All regressions control for ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects.

4.2. *Conditional random assignment*

Unlike most research settings, we know the exact procedures for individual assignment into peer groups. Since the assignment is quasi-random, this aids in estimating unbiased causal effects of rank because trainees are unable to self-select into flights where they would enjoy a more advantageous position. In our setting, trainees are placed into gender segregated flights as they arrive to BMT on a bus in a first-come, first-served basis, irregardless of ability level. Chesney et al. (2024) provides a detailed analysis demonstrating that conditioning on a small set of observables, such as their weekly cohort, MEPS, occupation, and gender, sufficiently describes the assignment process. For this setting, it is adequate to show that rank is uncorrelated with observable characteristics; therefore, by controlling for MEPs, occupation, and gender, assignment to flight, and rank itself, is exogeneous to the member.

We test our identifying assumptions by running the following balance test:

$$R_{ifc} = \alpha + \beta X_i + f(Ability_i) + \gamma_i + \phi_{fc} + \varepsilon_{ifc} \tag{2}$$

where we report the coefficients of X_i for individual i's cognitive and physical ranking in Table 4. Of the 18 estimated coefficients, only three, are statistically significant at the 5-percent level or below. Importantly, the magnitudes of all of the coefficients are small in magnitude, with no pattern in the direction of the signs. Hence, these results provide further suggestive evidence of the exogeneity of rank assignments in our setting.

5. Results

5.1. *Cognitive rank effects*

We report our main results of Eq. (1) in Table 5 for cognitive ordinal ranking effects. The first row of results shows average effects, while the remaining rows provide an interaction of ordinal ranking with gender. Overall, we find a one standard deviation increase in rank improves the probability of BMT completion and honor graduate

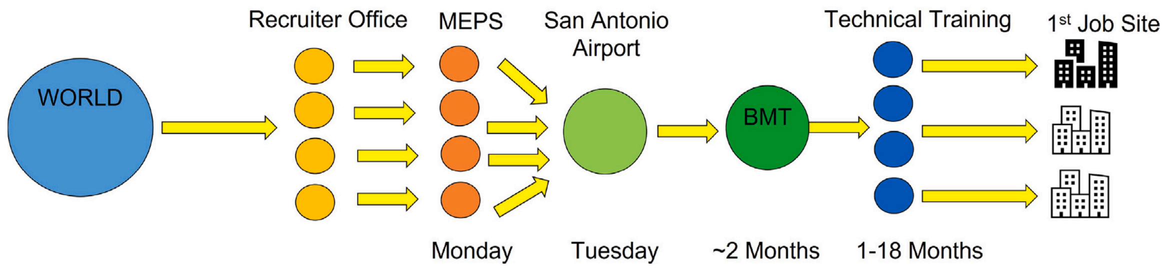

Fig. 10. Trainee arrival to/from basic military training.
The above picture illustrates how trainees arrive to BMT from their hometown, and depart for follow-on training and their first job site at a military installation.

Table 5
Cognitive rank outcomes.

	(1) BMT Completion	(2) Honor Graduate	(3) No Verbal Counseling	(4) No Disciplinary Action
Cognitive Ordinal Rank	0.032**	0.056***	0.006	0.036*
	(0.010)	(0.013)	(0.023)	(0.017)
N	215,132	215,132	159,693	159,693
r2	0.040	0.152	0.196	0.056
Cognitive Ordinal Rank	0.023*	0.051***	0.041	0.040*
	(0.010)	(0.014)	(0.024)	(0.017)
Rank × Female	0.018***	0.010	−0.090***	−0.010
	(0.005)	(0.006)	(0.010)	(0.007)
N	215,132	215,132	159,693	159,693
r2	0.040	0.152	0.196	0.056
Mean	0.939	0.091	0.663	0.867

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficient from the ordinal ranking in cognitive ability. Each regression controls for trainee background characteristics, ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 2 represent contemporaneous outcomes from Basic Military Training. Columns 3 and 4 represent disciplinary outcomes from follow-on technical training and the first 36 months at work. In the bottom half of the table, an interaction between rank and female is included.

status by a statistically significant 0.163 and 0.286 percentage points,[19] respectively. These models assume the rank effects are linear across own ability. To test this assumption, similar to Denning et al. (2023), Fig. 11 shows effects by ventile dummies of rank. These results indicate that the cognitive rank effects on BMT completion are largely linear. For example, moving from being ranked last to being ranked in the 25th percentile increases the probability of BMT completion by roughly 1.3 percentage points. Though sizeable, this effect is about one-third of that that found by Denning et al. (2023) for third grade retention. However, the effects for honor graduate status show substantial heterogeneity. Perhaps not too surprisingly, the effects are null for trainees in the bottom half of the rank distribution and are increasingly positive and significant for trainees in the top half of the rank distribution. For example, moving from being ranked in the 75th percentile to the top of the group increases the likelihood of becoming an "honor" graduate by approximately 3.3 percentage points. This effect is quite large given a 9.1 percent mean honor graduate rate.

Next, while rank has no discernible average effect on receiving "no verbal counseling" during follow-on technical training, we do find that a one standard deviation increase in rank increases the likelihood of "no disciplinary action" by 0.184 percentage points.[20] When examining results separately by gender, for women we find a stronger response of rank for completion and honor graduate status. However, we find weaker effects on no verbal counseling and no disciplinary action.

Additionally, we augment Eq. (1) to include interactions between ordinal rank and race, with results shown in Table 7. Overall, we find underrepresented minorities are *less* sensitive to rank compared to white trainees for short-term outcomes. This result is in contrast to Denning et al. (2023) who find that traditionally disadvantaged students in elementary school are *more* sensitive to rank.

5.2. Physical rank effects

For physical ranking effects, measured by initial scores on the fitness exam, we report our results of Eq. (1) in Table 6. Similar to Table 5, the first row presents average effects and the remaining rows provide an interaction of ordinal ranking with gender. Overall, physical rank shows no significant effects on BMT completion or verbal counseling. However, moving up in physical rank leads to positive and significant increases in the probability of becoming an honor graduate (0.277 p.p) and no disciplinary action (0.426 p.p.).[21] Next, to examine the linearity assumption, we graph the results in Fig. 11 by rank ventiles for BMT completion and honor graduate status. Results show physical rank has little impact on BMT graduation, save for the lowest ranked individuals. For honor graduate status, physical rank largely affects both the highest and lowest ranked individuals.

[19] These effect sizes are calculated by multiplying the rank coefficients in Table 5 by the standard deviation in the ordinal rank after controlling for ability and flight FE, which can be found in Table 2.

[20] In Appendix A1, we conduct robustness checks of our main specification accounting for missing data through attrition. Columns 1 and 2 estimate BMT outcomes while excluding non-active duty members. In the spirit of Horowitz and Manski (2000), Lee (2009), we conduct a bounding exercise for our long-run outcomes in columns 3 through 6 taking into consideration individuals who do not complete BMT. Overall, our findings are consistent with the main results. Importantly, the lower bounds of our estimated rank effects all fall within one-half of one standard error of our main estimates, indicating that attrition from the sample is not driving our estimated effects.

[21] Again, these effect sizes are calculated by multiplying the rank coefficient by the standard deviation in the ordinal rank after controlling for ability and flight FE.

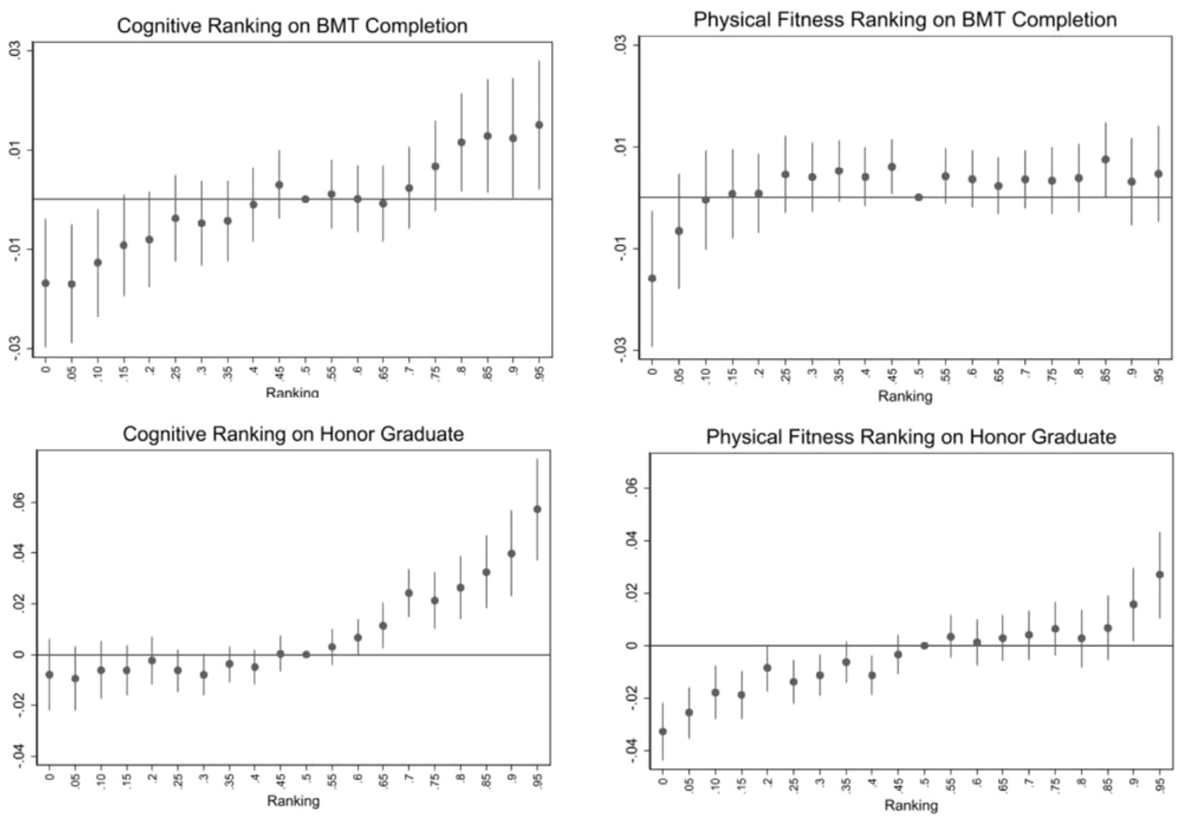

Fig. 11. Effect of cognitive and physical rank on BMT outcomes.
Note: The graphs above plot the coefficient for ventiles of ordinal cognitive and physical rank on short-run outcomes at BMT. The 45th to 50th percentile interval is the omitted category with 95% confidence intervals calculated using standard errors clustered at the cohort level. Each regression controls for trainee background characteristics, ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects.

Table 6
Physical rank outcomes.

	(1) BMT Completion	(2) Honor Graduate	(3) No Verbal Counseling	(4) No Disciplinary Action
Physical Ordinal Rank	0.011	0.039***	0.016	0.060***
	(0.008)	(0.009)	(0.016)	(0.013)
N	211,930	211,930	159,122	159,122
r2	0.061	0.116	0.185	0.059
Physical Ordinal Rank	−0.004	0.043***	0.017	0.070***
	(0.008)	(0.009)	(0.017)	(0.013)
Rank × Female	0.032***	−0.009	−0.002	−0.022**
	(0.005)	(0.005)	(0.010)	(0.007)
N	211,930	211,930	159,122	159,122
r2	0.062	0.116	0.185	0.059
Mean	0.949	0.092	0.663	0.867

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficient from the ordinal ranking in physical fitness. Each regression controls for trainee background characteristics, ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 2 represent contemporaneous outcomes from Basic Military Training. Columns 3 and 4 represent disciplinary outcomes from follow-on technical training and the first 36 months at work. In the bottom half of the table, an interaction between rank and female is included.

Interestingly, when estimating the effects interacted with gender, we find larger effects for women on BMT completion. For men, we find larger effects with honor graduate and no disciplinary action, compared to the full sample estimates. Finally, as done with our cognitive rank analysis, we augment Eq. (1) to include interactions between physical rank and race in Table 8. Again, we find underrepresented minorities are *less* sensitive to rank.

5.3. *Mechanisms*

5.3.1. *Environmental - leadership selection*

Elsner et al. (2021), Murphy and Weinhardt (2020) express how the “big-fish–little-pond effect” is a driving mechanism for how rank effects determine future outcomes. Specifically, Elsner et al. (2021) finds little evidence that effort based on ranking is a leading mechanism. In our

Table 7
Cognitive rank outcomes - by race.

	(1) BMT Completion	(2) Honor Graduate	(3) No Verbal Counseling	(4) No Disciplinary Action
Cognitive Ordinal Rank	0.036***	0.077***	0.014	0.037*
	(0.010)	(0.013)	(0.024)	(0.017)
Rank × Black	−0.008	−0.071***	−0.036**	0.002
	(0.005)	(0.006)	(0.011)	(0.009)
Rank × Hispanic	−0.012*	−0.025***	−0.000	−0.004
	(0.005)	(0.006)	(0.011)	(0.007)
Rank × Asian	−0.014*	−0.059***	−0.033*	−0.008
	(0.007)	(0.010)	(0.017)	(0.011)
N	215,132	215,132	159,693	159,693
r2	0.040	0.153	0.196	0.056
Mean	0.939	0.091	0.663	0.867

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficient from the ordinal ranking in cognitive ability and the interaction between ranking and race. Each regression controls for trainee background characteristics, ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 2 represent contemporaneous outcomes from Basic Military Training. Columns 3 and 4 represent disciplinary outcomes from follow-on technical training and the first 36 months at work.

Table 8
Physical rank outcomes - by race.

	(1) BMT Completion	(2) Honor Graduate	(3) No Verbal Counseling	(4) No Disciplinary Action
Physical Ordinal Rank	0.023**	0.075***	0.019	0.073***
	(0.009)	(0.009)	(0.017)	(0.013)
Rank × Black	−0.029***	−0.117***	−0.019	−0.040***
	(0.004)	(0.004)	(0.010)	(0.009)
Rank × Hispanic	−0.026***	−0.073***	0.010	−0.021**
	(0.005)	(0.005)	(0.010)	(0.008)
Rank × Asian	−0.046***	−0.017*	−0.021	−0.031**
	(0.007)	(0.009)	(0.016)	(0.011)
N	211,930	211,930	159,122	159,122
r2	0.062	0.119	0.185	0.059
Mean	0.939	0.091	0.663	0.867

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficient from the ordinal ranking in physical fitness and the interaction between ranking and race. Each regression controls for trainee background characteristics, ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 2 represent contemporaneous outcomes from Basic Military Training. Columns 3 and 4 represent disciplinary outcomes from follow-on technical training and the first 36 months at work.

Table 9
Leadership selection.

	(1) Element Leader	(2)	(3) Dorm Leader	(4)	(5) Any Leadership	(6)
Cognitive Ordinal Rank	−0.004		−0.004		−0.017	
	(0.011)		(0.006)		(0.021)	
Physical Ordinal Rank		0.069***		0.020***		0.108***
		(0.008)		(0.005)		(0.016)
N	215,132	211,930	215,132	211,930	215,132	211,930
r2	0.016	0.033	0.022	0.029	0.038	0.071
Cognitive Ordinal Rank	−0.016		−0.006		−0.033	
	(0.012)		(0.007)		(0.021)	
Cognitive Rank × Female	0.024***		0.005		0.031***	
	(0.005)		(0.003)		(0.009)	
Physical Ordinal Rank		0.060***		0.015**		0.078***
		(0.009)		(0.005)		(0.016)
Physical Rank × Female		0.018***		0.010***		0.063***
		(0.005)		(0.002)		(0.009)
N	215,132	211,930	215,132	211,930	215,132	211,930
r2	0.016	0.033	0.023	0.029	0.038	0.071
Mean	0.080	0.081	0.020	0.020	0.462	0.468

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficient from the ordinal ranking in cognitive and physical ability on the selection for leadership positions throughout BMT. The dataset reports 24 total possible leadership positions that a trainee could be selected into that range in responsibility. The most demanding is dorm leader, followed by element leader. Each regression controls for trainee background characteristics, ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 2 represent selection as an element leader. Columns 3 and 4 represent selection as a dorm leader. Columns 5 and 6 represent selection into any leadership position throughout BMT. In the bottom half of the table, an interaction between rank and gender is included.

research setting, we demonstrate how cognitive and physical ordinal ranking have differential impacts in both environmental (i.e. increased attention and resources from peers and instructors) and behavioral (i.e. individual effort due to a higher self-perception).

In the first part of this analysis, we examine how a higher cognitive and physical ranking impacts the likelihood of being selected for a leadership role during the early days of BMT. In our research setting, trainees are assigned various roles during BMT that range from a dorm leader with a very large amount of responsibility to being in charge of keeping the schedule moving. In total, there are 24 potential BMT leadership roles a trainee could be assigned. In total, we observe only 47% of the sample being assigned a leadership role during BMT. The most demanding role is being assigned "Dorm Leader" which represents 2% of our sample. The next most demanding job is "Element Leader" which represents 8% of our sample. We re-estimate our main model in Eq. (1) and report the results in Table 9. We show that cognitive ordinal ranking has almost no impact on being selected for any BMT leadership role. However, for women we find it has an outsized impact on selection for element leader and other leadership roles during BMT.

Additionally, we find physical ranking has a significantly large impact in the likelihood of being selected for a BMT leadership role across all the categories. These results suggest that instructors identify the highest ranking individuals early in BMT and provide them additional responsibilities. This environmental channel can lead to additional mentoring from BMT staff and attention from one's peers. This mechanism could arguably increase self-perception of one's abilities relative to peers, leading to potentially more effort exerted.

5.3.2. Behavioral - occupation choice

Another unique aspect of our research setting is that almost one-third of trainees that arrive to BMT during this time period do not have a preassigned post-BMT occupation in the Air Force. Instead, these individuals enlist in the USAF under the Aptitude Index (AI) program in which they are guaranteed a job under one of four specialities known as General, Administrative, Mechanical, and Electrical (National Academies of Sciences, Engineering, and Medicine, 2021). Individuals who meet the requirements and enlist under an aptitude area are then eligible for a list of occupations for which they qualify. As such, trainees arriving at BMT under these contracts receive a list of available jobs in their respective area around the second week of training. During the following weeks of training, these individuals then meet with a occupational specialists and further discuss the occupations they are qualified for after BMT graduation and compile a prioritized list of their preferences. Prior to graduating BMT, trainees are matched with an occupation based on their preference ranking, ASVAB scores, and other information about the trainee prior to arrival at BMT. Importantly, occupational match does not take into account actual performance during BMT.

Although we do not have access to each trainees prioritized list of occupations, we ultimately observe in our data the occupation with which they are matched. Additionally, we are able to assess the *skill level* of each matched job using the average ASVAB scores of individuals who were previously assigned these occupations. As such, we examine whether trainees' ordinal ranking in BMT impacts their assigned occupation by studying how it affects the ASVAB scores in their eventual job match. Because actual performance during BMT is not factored into job match, we believe the primary mechanism in how the trainee decides to specialize in a particular job is based on their perceived ability, through ordinal ranking within the flight.

We estimate these effects by again using Eq. (1), but instead normalize the ordinal rank measure to have a mean of 0 and a standard deviation of 1 to improve interpretability. Results in Table 10 show a one standard deviation increase in cognitive ordinal rank leads to a matched job with a 6% higher standard deviation in average AFQT score. We find similar sized effects across all four major components of the ASVAB. With respect to physical ordinal ranking, we estimate small, negative, and statistically insignificant effects on career specialization. These findings are counter to previous literature that show peers have little impact on job selection (Jones and Kofoed, 2020). This discrepancy is perhaps due to differences in settings and the underlying mechanisms. First, Air Force BMT trainees spend nearly every minute of the day interacting with their peers, with the explicit goal of training for their future job in the Air Force. Second, we specifically analyze ordinal ranking effects in which having a higher position may provide greater confidence to select, and subsequently be matched with, higher skilled jobs. Together, these results imply that decisions to specialize are highly dependent on social interactions, as previously found by Elsner et al. (2021) for academic fields of study. These findings complement the ordinal ranking literature by showing how higher-self perception is at play by individuals with higher ordinal rank.

5.4. Rank interactions

Given we find positive and significant rank effects for both our cognitive and physical measures, a natural question is how and whether these rank effects interact with one another? Understanding these dynamics may have important policy implications and, to our knowledge, rank interactions have not been studied in the previous literature. We begin this analysis by examining how rank effects interact with one another by estimating the following model for the full sample as well as separately for both men and women:

$$Y_{ifc} = \alpha + \beta_1 R^{Cog}_{ifc} + \beta_2 R^{Fit}_{ifc} + \beta_3 R^{Cog}_{ifc} * R^{Fit}_{ifc} + f(Ability_i) + X_i + \gamma_i + \phi_{fc} + \varepsilon_{ifc} \tag{3}$$

where R^{Cog}_{ifc} is cognitive ranking measured by AFQT scores and R^{Fit}_{ifc} is physical ranking measured by the initial fitness test at BMT. The parameter interest β_3 captures the interaction of these ranking measures. This specification controls for ventiles of ability for both cognitive and physical test scores in $f(Ability_i)$. We report the coefficients for β_1, β_2, and β_3 in Table 11.

For graduation, we find a positive coefficient for β_1 and β_2, but a negative and statistically significant value for the interaction term, β_3. This suggests that cognitive and physical fitness rankings act against one another for an individual completing the program. For example, column one demonstrates that a one standard deviation increase in cognitive ability can be cancelled out by a two standard deviation increase in physical ability. We see a similar pattern for men in column three and women in column five. For honor graduate status, we find a large positive and statistically significant value for the interactions term, β_3. These results are similar for both men and women as shown in columns four and six. These results suggest that cognitive and physical fitness rankings are multiplicative for obtaining honor graduate status from BMT.

These empirical finding on multiple ranking effects playing a dominant role in determining honor graduate status, yet working against each other in program completion, initially suggests ambiguous policy implications. For example, a policy maker who seeks to improve graduation rates for specific trainees would consider moving individuals with lower academic ability to a peer group where they enjoy a higher cognitive ranking. However, if the policy maker does not consider their new physical ordinal rank, they risk harming that individual's graduation prospects. Likewise, if they move into a group where they improve in rank on both categories, a policy maker inadvertently increases the likelihood of making an honor graduate.

6. Discussion

The heterogeneous response to ordinal rank effects, and the interactive effects between different rankings, motivates a deeper discussion on the policy implications. Previous literature by Bhattacharya (2009) laid the groundwork to exploit heterogeneous peer effects to

Table 10
Occupation selection.

	(1) AFQT	(2) Mechanical	(3) Administrative	(4) General	(5) Electronics
SD in Cognitive Rank	0.062**	0.047*	0.058*	0.066**	0.054**
	(0.024)	(0.020)	(0.024)	(0.023)	(0.020)
N	53,881	53,881	53,881	53,881	53,881
r2	0.377	0.522	0.367	0.381	0.540
SD in Physical Rank	−0.010	−0.017	−0.008	−0.011	−0.016
	(0.017)	(0.015)	(0.017)	(0.017)	(0.014)
N	53,881	53,881	53,881	53,881	53,881
r2	0.342	0.501	0.331	0.348	0.517

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficients from the rank effect. The rank variable is standardized to have a mean of 0 and a standard deviation of 1. The outcome of interest indicates the average test score from the ASVAB of a matched occupation during BMT. The results are reported for only trainees that arrive to training without a specific occupation assigned. Each regression controls for trainee background characteristics, ventile of achievement for both physical and cognitive test scores, aptitude area, MEPs station, and cohort-flight fixed effects. Column 1 reports average AFQT score of assigned occupation. Columns 2 through 5 report average Mechanical, Administrative, General, and Electronic scores from the ASVAB of the assigned occupation.

Table 11
Rank interactions.

	(1) Full Sample	(2)	(3) Men	(4)	(5) Women	(6)
	BMT Completion	Honor Graduate	BMT Completion	Honor Graduate	BMT Completion	Honor Graduate
Cognitive Ordinal Rank	0.041***	−0.180***	0.046***	−0.177***	0.048*	−0.187***
	(0.010)	(0.014)	(0.011)	(0.017)	(0.024)	(0.022)
Physical Ordinal Rank	0.019*	−0.199***	0.001	−0.195***	0.034	−0.153***
	(0.009)	(0.009)	(0.009)	(0.012)	(0.020)	(0.016)
Cognitive x Physical Rank	−0.017**	0.488***	−0.017**	0.505***	−0.017	0.429***
	(0.006)	(0.010)	(0.006)	(0.010)	(0.013)	(0.015)
N	211,930	211,930	163,797	163,797	47,905	47,905
r2	0.062	0.207	0.051	0.206	0.071	0.211
Mean	0.949	0.092	0.955	0.098	0.936	0.071

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (3) and report the coefficients from cognitive and physical ordinal ranking and the interaction. Each regression controls for trainee background characteristics, ventile of achievement for both physical and cognitive test scores, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 2 represent contemporaneous outcomes from Basic Military Training for all observations. Columns 3 and 4 report only men. Columns 5 and 6 report only women.

Table 12
Honor graduate on long-term outcomes.

	(1) No Failed Exam	(2) No Block Failure	(3) No Counseling	(4) No Disciplinary	(5) Career Change
Honor Graduate at BMT	0.076***	0.012***	0.065***	0.030***	0.004*
	(0.004)	(0.002)	(0.004)	(0.003)	(0.002)
N	159,122	159,122	159,122	159,122	156,494
r2	0.158	0.354	0.198	0.060	0.102
Mean	0.685	0.849	0.634	0.655	0.033
Men Only	0.075***	0.010***	0.064***	0.031***	0.004*
	(0.004)	(0.002)	(0.005)	(0.003)	(0.002)
N	126,608	126,608	126,608	126,608	124,440
r2	0.144	0.319	0.193	0.059	0.082
Mean	0.708	0.877	0.645	0.667	0.030
Women Only	0.073***	0.018**	0.066***	0.021***	0.005
	(0.009)	(0.007)	(0.010)	(0.006)	(0.005)
N	32,514	32,514	32,514	32,514	32,054
r2	0.189	0.393	0.224	0.064	0.157
Mean	0.610	0.758	0.599	0.614	0.044

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (4) and report the coefficients from honor graduate status. Each regression controls for trainee background characteristics, ventile of achievement for both physical and cognitive test scores, occupation, MEPs station, and cohort-flight fixed effects. Columns 1-3 represent outcomes in follow-on training. Columns 4-5 report disciplinary and career change indicators during an individual's first 36 months in the organization.

achieve pareto improving outcomes. More recent work by Opper et al. (2022) further builds on this idea of intentional assignment of trainees in job assignment programs to improve the overall mean. However, there has been little discussion on whether an assignment rule exists that improves individual outcome using rank effects. In our research setting, policy makers could tailor peer groups to support specific subgroups and leverage the heterogeneous responses to improve individual outcomes. One such mechanism is through honor graduate status.

Table A.1
Active duty sample and attrition bounding exercise.

	(1) BMT Completion	(2) Honor Graduate	(3) No Verbal Counseling	(4) No Verbal Counseling	(5) No Discipline	(6) No Discipline
			(1 - Attrit)	(0 - Attrit)	(1 - Attrit)	(0 - Attrit)
Cognitive Ordinal Rank	0.036**	0.078***	−0.007	0.030	0.030	0.066***
	(0.012)	(0.014)	(0.023)	(0.023)	(0.016)	(0.018)
	[0.032**]	[0.056***]	[0.006]	[0.006]	[0.036*]	[0.036*]
N	170,698	170,698	170,698	170,698	170,698	170,698
r2	0.045	0.147	0.180	0.168	0.053	0.051
Physical Ordinal Rank	0.009	0.047***	0.013	0.022	0.054***	0.063***
	(0.009)	(0.010)	(0.016)	(0.016)	(0.012)	(0.014)
	[0.011]	[0.039***]	[0.016]	[0.016]	[0.060*]	[0.060*]
N	168,051	168,051	168,051	168,051	168,051	168,051
r2	0.069	0.120	0.173	0.167	0.054	0.068

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficients from cognitive and physical ordinal ranking. These regressions only include active duty trainees which compose 79% of the total sample. Each regression controls for trainee background characteristics, ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 2 report contemporaneous outcomes from Basic Military Training. Columns 3 and 4 report disciplinary outcomes from follow-on technical training. Columns 5-6 report disciplinary outcomes during the first 36 months in the organization. In Columns 3 and 5, trainees that do not complete BMT are included and the outcome of interest is coded as 1. In Columns 4 and 6, the outcomes are given a value of 0. These columns provide a bound on attrition bias in which our main results reported in Tables 5 and 6 with our original point estimates displayed in brackets below. Importantly, the lower bounds of our estimated rank effects all fall within one-half of one standard error of our main estimates, indicating that attrition from the sample is not driving our estimated effects.

Table A.2
Missing physical fitness bounding exercise.

	(1)	(2)	(3)	(4)	(5)	(6)
	BMT Completion			Honor Graduate		
	Low Value	Random Value	High Value	Low Value	Random Value	High Value
Physical Ordinal Rank	−0.022*	0.003	0.082***	0.041***	0.045***	0.051***
	(0.010)	(0.008)	(0.013)	(0.009)	(0.009)	(0.009)
		[0.011]			[0.039***]	
N	215,132	215,132	215,132	215,132	215,132	215,132
r2	0.118	0.066	0.097	0.117	0.116	0.113
	No Verbal Counseling			No Disciplinary Action		
Physical Ordinal Rank	0.011	0.018	0.017	0.061***	0.063***	0.068***
	(0.016)	(0.016)	(0.016)	(0.013)	(0.013)	(0.012)
		[0.016]			[0.060***]	
N	159,693	159,693	159,693	159,693	159,693	159,693
r2	0.185	0.185	0.185	0.059	0.059	0.059

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficient from the ordinal ranking in physical fitness. Each regression controls for trainee background characteristics, ventile of achievement, occupation, MEPs station, and cohort-flight fixed effects. This table presents the results of a bounding exercise for the 3202 trainees that do not take an initial fitness exam upon arrival at BMT for an undisclosed reason. Of those 3202 trainees, only 697 (21.77%) complete BMT and none of them become honor graduates, 448 (13%) receive no counseling and 751 (23.45%) no disciplinary in follow on training. Columns 1 and 3 replace the 3202 missing values with the lowest score on the fitness exam. Columns 2 and 5 replace the missing values with a random value. Columns 3 and 6 replace the missing value with the highest score on the fitness exam. The actual estimates from Table 6 are reported in brackets in columns 2 and 5.

Our prior analysis demonstrates that ordinal ranking increases the likelihood of becoming an honor graduate. BMT designates no more than 10 percent of its trainees for this award. This is calculated using performance in final academics and fitness exams, daily inspections, and serving in leadership roles throughout training. Although not causal, we estimate the following regression to describe the relationship between honor graduate status on long-term outcomes:

$$Y_i = \alpha + \beta Honor + f(Ability_i) + X_i + \gamma_i + \phi_{fc} + \varepsilon_{ifc} \tag{4}$$

where Y_i are a set of outcomes post-BMT, $Honor$ is an indicator of whether individual i is an honor graduate from BMT, $f(Ability_i)$ controls for academic and physical test scores, X_i and γ_i controls for personal characteristics, occupation, and MEPs, and ϕ_{fc} controls for flight-cohort fixed effects. Results are listed in Table 12. Across all outcomes for both men and women, honor graduate status has a strong positive association with improved academic performance, avoiding misconduct, and retraining into a new career-field after 24 months in the organization. As a result, technical training programs that seek to improve the outcomes of certain trainees that exhibit higher disciplinary or washout rates could use the initial assignment at BMT to *nudge* these trainees favorably. However appealing, previous research shows that using peer effects estimates to optimally assign individuals to groups is not as straightforward as it may seem (Carrell et al., 2013). Likewise, our results showing differing interaction effects between cognitive and physical rank across outcomes, suggests this optimization problem is even more complex. That is, policy makers wishing to use ordinal rank to form a deliberate assignment process will need take into account the multidimensional response to changes in rank. We leave this for future research.

7. Conclusion

In this paper, we analyzed trainees' ordinal rank across cognitive and physical ability during an initial job training program for the USAF. Rank is measured within peer groups that are assigned in an almost-as-good as random fashion. We find cognitive ordinal ranking, measured by predetermined scores on the AFQT, to have a meaningful impact on completing BMT and becoming an honor graduate. Rank additionally

Table A.3
Cognitive rank outcomes.

	(1) BMT Completion	(2)	(3)	(4) Honor Graduate	(5)	(6)
	Ventile Bins	2 pp Bins	Additional Controls	Ventile Bins	2 pp Bins	Additional Controls
Cognitive Ordinal Rank	0.032**	0.029**	0.032**	0.056***	0.033*	0.048***
	(0.010)	(0.010)	(0.010)	(0.013)	(0.014)	(0.013)
BMI Ordinal Rank			0.007			−0.014
			(0.010)			(0.012)
Height Ordinal Rank			0.004**			0.004**
			(0.002)			(0.002)
N	215,132	215,132	205,378	215,132	215,132	205,378
r2	0.040	0.040	0.062	0.152	0.153	0.187
	No Verbal Counseling			No Disciplinary Action		
Cognitive Ordinal Rank	0.006	−0.005	0.002	0.036*	0.035*	0.028
	(0.023)	(0.025)	(0.023)	(0.017)	(0.018)	(0.017)
BMI Ordinal Rank			−0.027			−0.010
			(0.024)			(0.017)
Height Ordinal Rank			−0.002			−0.008
			(0.004)			(0.003)
N	215,132	215,132	205,378	215,132	215,132	205,378
r2	0.196	0.196	0.197	0.056	0.056	0.062

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficient from the ordinal ranking in cognitive ability. Each regression controls for trainee background characteristics, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 4 control for ability using ventiles of achievement. Columns 2 and 5 control for ability using 2 percentage points of achievement. Columns 3 and 6 control for ability using ventile of ability and add the additional controls of BMI and Height Ordinal ranking.

Table A.4
Physical rank outcomes.

	(1) BMT Completion	(2)	(3)	(4) Honor Graduate	(5)	(6)
	Ventile Bins	2 pp Bins	Additional Controls	Ventile Bins	2 pp Bins	Additional Controls
Physical Ordinal Rank	0.011	0.004	0.012	0.039***	0.014	0.040***
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
BMI Ordinal Rank			0.007			−0.010
			(0.010)			(0.012)
Height Ordinal Rank			0.005**			0.017***
			(0.002)			(0.002)
N	211,930	211,930	205,378	211,930	211,930	205,378
r2	0.061	0.062	0.062	0.116	0.117	0.118
	No Verbal Counseling			No Disciplinary Action		
Physical Ordinal Rank	0.016	0.015	0.017	0.060***	0.050***	0.058***
	(0.016)	(0.017)	(0.017)	(0.013)	(0.013)	(0.013)
BMI Ordinal Rank			−0.024			−0.010
			(0.024)			(0.017)
Height Ordinal Rank			0.005			−0.006
			(0.004)			(0.003)
N	211,930	211,930	205,378	211,930	211,930	205,378
r2	0.185	0.185	0.186	0.059	0.060	0.062

Robust standard errors are clustered at the cohort level in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Note: Results from the above table are from Eq. (1) and report the coefficient from the ordinal ranking in cognitive ability. Each regression controls for trainee background characteristics, occupation, MEPs station, and cohort-flight fixed effects. Columns 1 and 4 control for ability using ventiles of achievement. Columns 2 and 5 control for ability using 2 percentage points of achievement. Columns 3 and 6 control for ability using ventile of ability and add the additional controls of BMI and Height Ordinal ranking.

impacts decisions to specialize in a more technical occupation. Physical ordinal ranking, measured by initial fitness scores taken during the first week at BMT, also affects early training outcomes. Both sets of ranking effects impact follow-on disciplinary outcomes and vary by gender. The interaction between cognitive and physical ordinal ranking is found to have multiplicative effects on honor graduate status, which is associated with important follow-on outcomes. Our heterogeneity analysis also illustrates that certain occupations and subgroups respond more favorably to rank effects. These results can serve as an important policy insight for improving outcomes of at-risk individuals. Future research should investigate how cognitive ranking compares with other descriptions of human capital, such as non-cognitive and social measures.

Declaration of competing interest

We authored this paper in support of Alex Chesney's PhD dissertation at the University of California, Davis. Alex is an active duty U.S. military member; however, this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We do not have any other conflicts of interest to declare.

Data availability

The authors do not have permission to share data.

Appendix. Robustness checks and bounding exercises

See Tables A.1–A.4.

References

Air Force Instruction 36-2907, 2014. Unfavorable information file (UIF) program. https://www.mcmilitarylaw.com/documents/afi36-2907.pdf.

Air Force Personnel Center, 2021. AIR FORCE ENLISTED CLASSIFICATION DIRECTORY (AFECD): The official guide to the air force enlisted classification codes. http://www.milvet.state.pa.us/DMVA/Docs_PNG/hro/AFECD.pdf.

Angrist, J.D., 1990. Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records. Am. Econ. Rev. 80 (3), 313–336.

Angrist, J.D., 1998. Estimating the labor market impact of voluntary military service using social security data on military applicants. Econometrica 66 (2), 249–288.

Antonovics, K., Golan, L., 2012. Experimentation and job choice. J. Labor Econ. 30 (2), 333–366.

Bertoni, M., Nisticò, R., 2023. Ordinal rank and the structure of ability peer effects. J. Public Econ. 217, 104797.

Bhattacharya, D., 2009. Inferring optimal peer assignment from experimental data. J. Amer. Statist. Assoc. 104 (486), 486–500.

Bloem, M., 2015. Becoming an MTI: What it takes. In: 149 Fighter Wing. SAF Public Affairs, https://www.149fw.ang.af.mil/News/Features/Display/Article/865310/becoming-an-mti-what-it-takes/.

Booij, A.S., Leuven, E., Oosterbeek, H., 2016. Ability peer effects in university: Evidence from a randomized experiment. Rev. Econ. Stud. 84 (2), 547–578.

Brown, G.D.A., Gardner, J., Oswald, A.J., Qian, J., 2008. Does wage rank affect employees' well-being? Ind. Relat. A J. Econ. Soc. 47 (3), 355–389.

Card, D., Mas, A., Moretti, E., Saez, E., 2012. Inequality at work: The effect of peer salaries on job satisfaction. Amer. Econ. Rev. 102 (6), 2981–3003.

Carrell, S.E., Sacerdote, B.I., West, J.E., 2013. From natural variation to optimal policy? The importance of endogenous peer group formation. Econometrica 81 (3), 855–882.

Chanal, J.P., Marsh, H.W., Sarrazin, P.G., Bois, J.E., 2005. Big-fish-little-pond effects on gymnastics self-concept: Social comparison processes in a physical setting. J. Sport Exerc. Psychol. 27 (1), 53–70.

Chesney, A., Cullen, J.B., Dahl, G., 2024. Measuring effectiveness outside of the classroom: Evidence from basic military training.

Cicala, S., Fryer, R.G., Spenkuch, J.L., 2017. Self-selection and comparative advantage in social interactions. J. Eur. Econom. Assoc. 16 (4), 983–1020.

Cornelissen, T., Dustmann, C., Schönberg, U., 2017. Peer effects in the workplace. Amer. Econ. Rev. 107 (2), 425–456.

Council on Foreign Relations, 2020. Demographics of the U.S. military. https://www.cfr.org/backgrounder/demographics-us-military.

Cullen, J.B., Dahl, G., Thorpe, R., 2024. The costs of job mismatch: Evidence using quasi-random assignment in the air force.

Dahl, G.B., Løken, K.V., Mogstad, M., 2014. Peer effects in program participation. Amer. Econ. Rev. 104 (7), 2049–2074.

Delaney, J., Devereux, P., 2022. Rank Effects in Education: What Do We Know So Far? Technical Report, Institute of Labor Economics (IZA).

Deming, D.J., 2017. The growing importance of social skills in the labor market*. Q. J. Econ. 132 (4), 1593–1640.

Denning, J.T., Murphy, R., Weinhardt, F., 2023. Class rank and long-run outcomes. Rev. Econ. Stat. 1426–1441.

Duflo, E., Dupas, P., Kremer, M., 2011. Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. Amer. Econ. Rev. 101 (5), 1739–1774.

Edin, P.-A., Fredriksson, P., Nybom, M., Öckert, B., 2022. The rising return to noncognitive skill. Am. Econ. J. Appl. Econ. 14 (2), 78–100.

Elsner, B., Isphording, I.E., 2017. A big fish in a small pond: Ability rank and human capital investment. J. Labor Econ. 35 (3), 787–828.

Elsner, B., Isphording, I.E., 2018. Rank, sex, drugs, and crime. J. Hum. Resour. 53 (2), 356–381.

Elsner, B., Isphording, I.E., Zölitz, U., 2021. Achievement rank affects performance and major choices in college. Econ. J. 131 (640), 3182–3206.

Epple, D., Romano, R.E., 2011. Chapter 20 - peer effects in education: A survey of the theory and evidence. In: Benhabib, J., Bisin, A., Jackson, M.O. (Eds.), In: Handbook of Social Economics, vol. 1, North-Holland, pp. 1053–1163.

Gill, D., Kissová, Z., Lee, J., Prowse, V., 2019. First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision. Manage. Sci. 65 (2), 494–507.

Greenberg, K., Gudgeon, M., Isen, A., Miller, C., Patterson, R., 2022. Army service in the all-volunteer era. Working Paper, https://drive.google.com/file/d/1w53cPug29Mas_ERrUa-eusPOfUwnVg-y/view.

Heckman, J.J., Zhou, J., 2022. Measuring Knowledge Working Paper 29990, National Bureau of Economic Research.

Horowitz, J.L., Manski, C.F., 2000. Nonparametric analysis of randomized experiments with missing covariate and outcome data. J. Amer. Statist. Assoc. 95 (449), 77–84.

Hoxby, C.M., 2000. The effects of class size on student achievement: New evidence from population variation. Q. J. Econ. 115 (4), 1239–1285.

Johnson, J.F., Romay, S., Barron, L.G., 2020. Air Force Work Interest Navigator (AF-WIN) to improve person-job match: Development, validation, and initial implementation. Mil. Psychol. 32 (1), 111–126.

Jones, T.R., Kofoed, M.S., 2020. Do peers influence occupational preferences? Evidence from randomly-assigned peer groups at West Point. J. Public Econ. 184, 104154.

Kiessling, L., Norris, J., 2022. The long-run effects of peers on mental health. Econ. J. 133 (649), 281–322.

Kim, J., 2021. Assessing the importance of relative body weight in shaping life satisfaction: Evidence from 1 million Americans. Curr. Psychol.

Klausmann, T., Wagner, V., Zipperle, I., 2021. Rank response functions in an online learning environment. Econom. Lett. 209, 110096.

Lang, K., Manove, M., 2011. Education and labor market discrimination. Amer. Econ. Rev. 101 (4), 1467–1496.

Lee, D.S., 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. Rev. Econ. Stud. 76 (3), 1071–1102.

Lindo, J.M., Sanders, N.J., Oreopoulos, P., 2010. Ability, gender, and performance standards: Evidence from academic probation. Am. Econ. J. Appl. Econ. 2 (2), 95–117.

Marrone, J.V., 2020. Predicting 36-Month Attrition in the U.S. Military: A Comparison Across Service Branches. RAND Corporation, Santa Monica, CA.

Marsh, H.W., 1987. The big-fish-little-pond effect on academic self-concept. J. Educ. Psychol. 79 (3), 280.

Marsh, H.W., Parker, J.W., 1984. Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? J. Pers. Soc. Psychol. 47, 213–231.

Mas, A., Moretti, E., 2009. Peers at work. Amer. Econ. Rev. 99 (1), 112–145.

McClean, H., 2021. Discharged and discarded: The collateral consequences of a less-than-honorable military discharge. Columbia Law Rev. 121 (7), 2203–2268.

Merton, R.K., 1968. The matthew effect in science. Science 159 (3810), 56–63.

Murphy, R., Weinhardt, F., 2020. Top of the class: The importance of ordinal rank. Rev. Econ. Stud. 87 (6), 2777–2826.

National Academies of Sciences, Engineering, and Medicine, 2021. Strengthening U.S. Air Force Human Capital Management: A Flight Plan for 2020–2030. The National Academies Press, Washington, DC.

Neal, D.A., Johnson, W.R., 1996. The role of premarket factors in black-white wage differences. J. Polit. Econ. 104 (5), 869–895.

Opper, I.M., Baird, M.D., Engberg, J., 2022. Optimal allocation of seats in the presence of peer effects: Evidence from a job training program. J. Labor Econ..

Pagani, L., Comi, S., Origo, F., 2021. The effect of school rank on personality traits. J. Hum. Resour. 56 (4), 1187–1225.

Rosen, S., 1981. The economics of superstars. Am. Econ. Rev. 71 (5), 845–858.

Rothstein, J., Rouse, C.E., 2011. Constrained after college: Student loans and early-career occupational choices. J. Public Econ. 95 (1), 149–163.

Sacerdote, B., 2014. Experimental and quasi-experimental analysis of peer effects: Two steps forward? Annu. Rev. Econ. 6 (1), 253–272.

SAFPA, 2021. USAF Demographics. Air Force News Service, https://www.afpc.af.mil/About/Air-Force-Demographics/.

Schaeffer, K., 2021. The Changing Face of America's Veteran Population. Pew Research Center, https://www.pewresearch.org/fact-tank/2021/04/05/the-changing-face-of-americas-veteran-population/.

Trent, J.D., Barron, L.G., Rose, M.R., Carretta, T.R., 2020. Tailored adaptive personality assessment system (TAPAS) as an indicator for counterproductive work behavior: Comparing validity in applicant, honest, and directed faking conditions. Mil. Psychol. 32 (1), 51–59.

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. Science 185 (4157), 1124–1131.

U.S. Bureau of Labor Statistics, 2022a. College enrollment and work activity of recent high school and college graduates summary. https://www.bls.gov/news.release/hsgec.nr0.htm.

U.S. Bureau of Labor Statistics, 2022b. Occupational outlook handbook, military careers. https://www.bls.gov/ooh/military/military-careers.htm.

Weingarten, E., Bhatia, S., Mellers, B., 2018. Multiple goals as reference points: One failure makes everything else feel worse. Manage. Sci. 65.