AFRL-SA-WP-TR-2015-0008

Multidimensional Aptitude Battery-Second Edition Intelligence Testing of Remotely Piloted Aircraft Training Candidates Compared with Manned Airframe Training Candidates

Julie Swearingen, PhD.[1]; Wayne Chappelle, PsyD, ABPP[2]; Tanya Goodman, MS[1]; William Thompson, MA[1]

[1]Neurostat Analytical Solutions, San Antonio, TX; [2]U.S. Air Force School of Aerospace Medicine, Aerospace Medicine Dept., Wright-Patterson AFB, OH

March 2015

Final Report
for September 2013 to December 2014

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

STINFO COPY

Air Force Research Laboratory
711th Human Performance Wing
U.S. Air Force School of Aerospace Medicine
Aeromedical Research Department
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

Qualified requestors may obtain copies of this report from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-SA-WP-TR-2015-0008 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNATURE//	//SIGNATURE//
LT COL SUSAN DUKES	DR. RICHARD A. HERSACK
Chief, Aircrew Selection & Performance Res	Chair, Aeromedical Research Department

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
1 Mar 2015

2. REPORT TYPE
Final Technical Report

3. DATES COVERED *(From – To)*
September 2013 – December 2014

4. TITLE AND SUBTITLE
Multidimensional Aptitude Battery-Second Edition Intelligence Testing of Remotely Piloted Aircraft Training Candidates Compared with Manned Airframe Training Candidates

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
Julie Swearingen, Wayne Chappelle, Tanya Goodman, William Thompson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
USAF School of Aerospace Medicine
Aeromedical Research Department/FHC
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL-SA-WP-TR-2015-0008

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The advancement of aviation drone technology has led to significant developments and improvements in the capabilities of military remotely piloted aircraft (RPAs, also known as drones). The prolific demand for RPA missions has led to an ever-increasing need for RPA pilots and to the development of a U.S. Air Force (USAF) RPA pilot career field. To date, there are limited objective data published on personnel who desire and self-select to USAF weapons bearing RPA pilots. This study evaluated the pre-training standardized general intelligence testing for three groups of pilot training candidates: Group 1 - newly commissioned officers who volunteered to become RPA pilots; Group 2 - pilot training candidates who completed undergraduate pilot training for a manned airframe but were reassigned to RPA pilot training; and Group 3 - training candidates who completed undergraduate pilot training and were assigned additional manned airframes. General intelligence testing consisted of the Multidimensional Aptitude Battery-Second Edition taken during initial medical flight screening prior to pilot training. Overall, the results of the study reveal the intellectual ability of those who are motivated and self-selected to pursue a career in RPA pilot training is very similar to those who are motivated to pursue manned airframe pilot training. However, performance in any demanding career field is more than a function of intellectual ability, and pilots are no exception. All three groups scored significantly higher than same-age peers in the general population on intellectual ability, but that does not mean that small differences in intellectual ability within these pre-screened, high-functioning populations will be insignificant when subsequently predicting performance and training success.

15. SUBJECT TERMS
Remotely piloted aircraft, RPA, pilot training, MAB-II, medical flight screening, intellectual ability, FSIQ, PIQ, VIQ

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Wayne Chappelle, PsyD
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	SAR	22	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This page intentionally left blank.

TABLE OF CONTENTS

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

1.0 EXECUTIVE SUMMARY

The advancement of aviation drone technology has led to significant developments and improvements in the capabilities of military remotely pilot aircraft (RPAs, also known as drones). Although the Department of Defense has a wide range of RPAs, the MQ-1 Predator and MQ-9 Reaper have emerged as the primary unmanned assets for acquiring intelligence and conducting surveillance and reconnaissance while carrying out close air support and sniper weapons strikes. The prolific demand for RPA missions has led to an ever-increasing need for RPA pilots and to the development of a U.S. Air Force (USAF) RPA pilot career field. Effective recruitment into this new career field is critical to meeting this demand, but has proven to be much more challenging than recruitment into traditional manned pilot airframes. To date, there are limited objective data published on personnel who desire and self-select to become USAF military-weapons-bearing RPA pilots.

This study evaluated the pre-training standardized general intelligence testing for three groups of pilot training candidates: Group 1, newly commissioned officers who volunteered to become RPA pilots upon entry in the USAF (n = 411); Group 2, pilot training candidates who completed undergraduate pilot training (UPT) for a manned airframe but were reassigned to RPA pilot training due to the need to fill RPA pilot vacancies (n = 36); and Group 3, training candidates who completed UPT and were assigned additional manned airframes, such as a fighter/bomber, surveillance/reconnaissance, tanker/transporter, or helicopter (n = 7,248). General intelligence testing consisted of the Multidimensional Aptitude Battery-Second Edition taken during initial medical flight screening prior to pilot training.

The results of the study indicated the average full scale intelligence quotient (IQ) and performance IQ scores for all three groups were in the superior range of intellectual ability. Average verbal IQ (VIQ) scores for all three groups were in the high average range, although modest differences were noted among the groups for VIQ, with self-selected RPA pilot trainees scoring slightly lower than both UPT trainees who cross-trained into RPAs and UPT trainees who went on to fly manned aircraft. A few mean differences were found for three of the VIQ subtests (comprehension, information, and arithmetic) and one performance IQ subtest (picture completion), which are discussed at length.

Overall, the results of the study reveal the intellectual ability of those who are motivated and self-selected to pursue a career in RPA pilot training is very similar to those who are motivated to pursue manned airframe pilot training. However, performance in any demanding career field is more than a function of intellectual ability, and aviators are no exception. All three of these aviator groups scored significantly higher than same-age peers on intellectual ability, but that does not mean that small differences in intellectual ability within these pre-screened, high-functioning populations will be insignificant in predicting performance and training success. Future research should focus on further investigating these modest differences in intellectual ability within the context of personality, motivation, and other cognitive skills and abilities which, when taken together, could help successfully recruit and retain our next generation of RPA pilots.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

2.0 INTRODUCTION

Remotely piloted aircraft (RPA) are part of a rapidly expanding area of U.S. Air Force (USAF) capabilities and missions within the continental United States and across the globe. In particular, the MQ-1 Predator and MQ-9 Reaper have emerged as dominant weapons-bearing RPAs that are considered critical to aerial intelligence, surveillance, reconnaissance, as well as close air support and sniper operations in various regions across the globe. USAF aviators stationed within the United States operate and control these airframes within theaters of conflict (i.e., Iraq and Afghanistan) and areas of national interest (i.e., Africa), which are often several thousand miles away [1,2].

Military flying in support of combat and/or humanitarian missions is an extraordinary profession requiring a special set of traits and talents. Many perceive that those who desire to become military pilots possess superior levels of intelligence, dexterity, coordination, and reflexes that are combined with a strong motivation to fly. This particular picture of a pilot has been portrayed in novels as well as films (e.g., *The Right Stuff* and *Top Gun*) and is a common perception among military leadership and civilians. However, having an accurate assessment of the cognitive aptitudes of USAF pilots is important to aeromedical providers tasked with evaluating rated pilots and training applicants and making decisions about whether such persons are aeromedically suitable to pursue such a challenging and high-risk occupation. Our current understanding of the psychological attributes considered critical to performance is based upon pilots in manned airframes (e.g., fighter/bomber, tanker/transporter, and surveillance/reconnaissance). Our understanding of the traits and attributes affecting performance among pilots of unmanned airframes is very limited.

A meta-analysis of military pilot selection literature over the past 20 years concluded that inherent cognitive aptitudes relevant to pilot performance include high levels of intelligence, dexterity, visual-spatial abilities, memory, attention/concentration, psychomotor reaction time, as well as speed and information processing [3]. Several studies assessing intelligence and cognitive aptitudes of USAF pilots of manned airframes [4-8] confirm their scores are typically in the superior range of functioning with significantly less variability when compared with peers in the general population. The finding that USAF pilots have a high level of cognitive aptitude is not surprising, given such aptitude is one of the strongest predictors of job performance in general [9,10], as well as pilot training [11,12]. Based upon the body of empirical findings, it stands to reason that high levels of intelligence and inherent cognitive aptitudes are critical to training and adapting to the operational demands of military flying.

However, the literature on cognitive aptitudes specific to the performance of RPA pilots is limited. A comprehensive review of the basic knowledge, skills, and abilities of RPA pilots in general (civilian and military) by Pavlas et al. alluded to several cognitive attributes as key to performance, including situational awareness, vigilance, spatial analyses (i.e., ability to mentally manipulate two-dimensional objects into a three-dimensional mental image), reasoning, speed of information processing, as well as visual tracking, searching, and scanning [13]. The results of their review were similar to other studies that assessed the job tasks and skills required for military-specific RPAs such as the Pioneer [14,15], Global Hawk [16], and Predator [17]. Cognitive aptitudes that appear common to most major Predator job accomplishments include high levels of situational awareness, vigilance, spatial analyses and reasoning, speed of information processing, visual tracking, searching, and scanning, as well as complex and divided attention. A comprehensive tasks analysis focusing on cognitive aptitudes critical to Predator

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

pilots completed by Bailey identified the following as critical to performance: perceptual reasoning and processing, short-term memory, spatial reasoning, symbolic reasoning, central information processing, psychomotor dexterity, and reaction time [18]. Bailey reasoned that cognitive aptitudes contribute to about two-thirds of the factors associated with Predator pilot job training and success. Lastly, USAF subject matter experts (e.g., RPA squadron commanders, training instructors, and trained pilots) reported high levels of cognitive proficiency, visual perception, attention, spatial processing, memory, reasoning, and psychomotor reacting time as critical to successfully completing training and adapting to operational rigors [19].

To date, there is only one published report in the general literature empirically assessing the general intelligence and neuropsychological aptitudes of USAF RPA pilot training candidates. Chappelle et al. obtained comprehensive computer-based intelligence testing (Multidimensional Aptitude Battery-Second Edition [MAB-II]) and neuropsychological screening (MicroCog) on USAF MQ-1 Predator nonrated pilot training candidates who passed the initial RPA flying screening course (n=108), nonrated training candidates who failed the training course (n=52), as well as USAF rated pilot training candidates who cross-trained to the MQ-1 Predator from manned airframes (n=157) [20]. The results of the study revealed nonrated pilot training candidates performed in the high average to superior range of functioning and nonrated pilot training candidates who passed training scored higher on measures of spatial analyses/reasoning, memory for novel spatial arrangements, general visual reasoning, visual construction, general executive reasoning, and general information processing accuracy when compared with nonrated pilot training candidates who failed training. Furthermore, nonrated pilot training candidates who passed training performed substantially higher on measures of spatial analyses/reasoning, memory for novel spatial arrangements, visual reasoning, general information processing accuracy, and cognitive proficiency (a combination and accuracy of speed of information processing) in comparison to rated pilots who cross-trained from a manned airframe. The results of the study provide insights into the aptitudes needed to adapt to the rigors of the training program, as well as the cognitive capabilities of those training candidates newly recruited for this career field. However, minimal studies exist comparing and contrasting RPA pilot training candidates with those candidates assigned to manned airframe pilot training.

Given the general lack of data regarding the breadth and level of cognitive abilities of those who volunteer and self-select to become RPA pilots in comparison to those who self-select to become manned airframe pilots, the purpose of the current study is to compare scores on the MAB-II for three pilot training candidate groups. The study evaluates newly commissioned officers who entered the USAF and volunteered to become RPA pilot candidates, RPA pilot training candidates who graduated pilot training for a manned airframe but were cross-trained due to the need to fill RPA pilot vacancies, and manned aircraft pilot training candidates who completed pilot training and were selected for upgrade specialty training in a manned airframe (such as fighter/bomber, tanker/transporter, or surveillance/reconnaissance) to assess for between-group differences in general intellectual ability among the separate groups of pilot trainees. The results and findings are likely to have implications for personnel selection, classification, and aeromedical evaluation processes.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

3.0 METHODS

3.1 Participants

This study evaluated participants who entered undergraduate pilot training (UPT) and MQ-1 RPA pilot training from 2008 to 2013. Figure 1 is a visual representation of accession sources for the three groups assessed in this study.

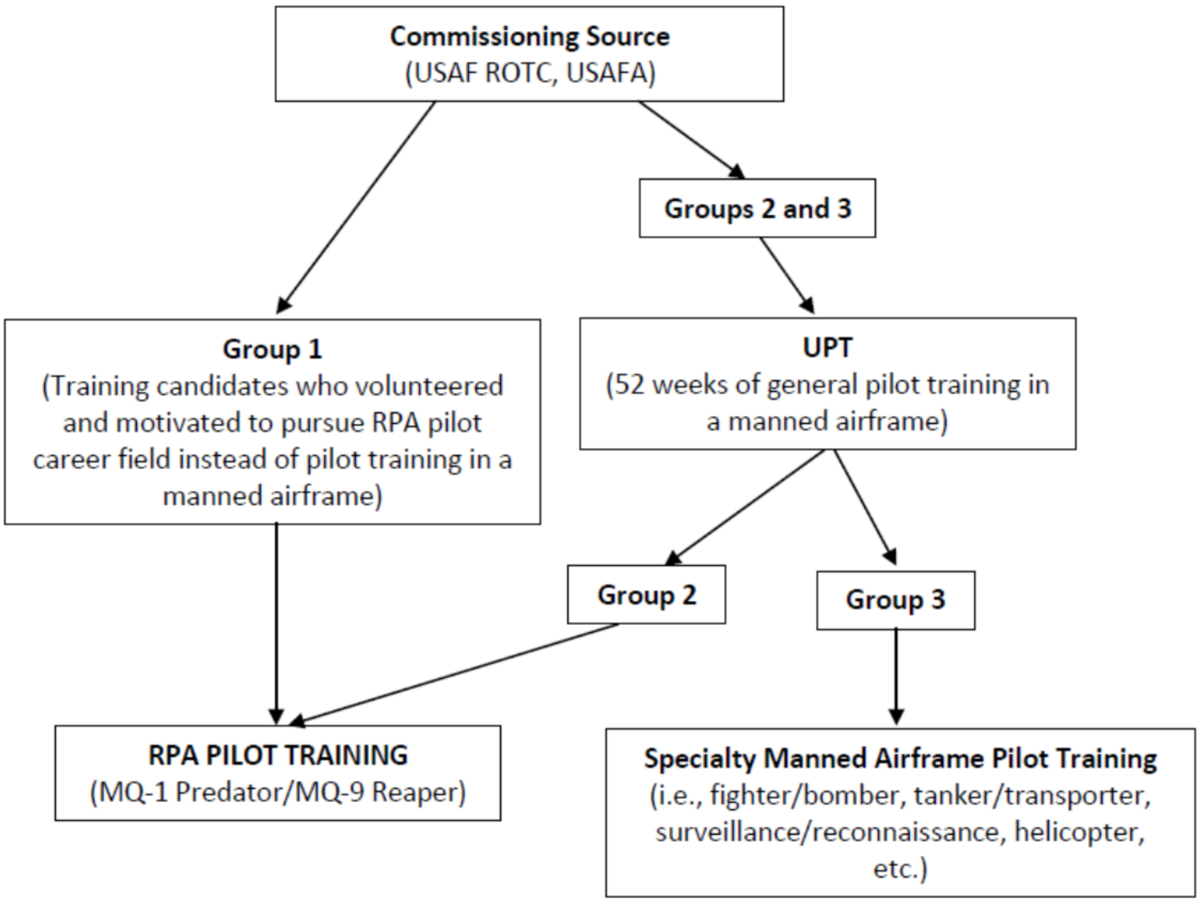

Figure 1. RPA pilot training candidate accession sources. ROTC = Reserve Officer Training Corps at civilian universities and colleges; USAFA = United States Air Force Academy.

3.1.1 Group 1: RPA-Direct Pilot Training Candidates. Group 1 consists of newly commissioned officers who volunteered to go directly into RPA MQ-1 Predator pilot training after completing all necessary application and selection requirements. Group 1 did not receive previous training to become pilots for manned military aircraft.

In total, 411 RPA-direct pilot training candidates were included in this study. This group consisted of 371 (90.28%) male and 34 (8.27%) female participants with an average age of 24.22 (standard deviation [*SD*] = 3.01) years. A total of 322 training candidates reported Caucasian (78.35%) as their ethnicity, followed by 30 Asian/Pacific Islander (7.30%), 20 Hispanic (6.08%), 17 African (4.14%), and three Indian (American)/Eskimo/Aleut participants (0.73%).

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

3.1.2 Group 2: RPA Pilot Training Candidates Who Completed UPT. Group 2 consists of RPA pilot training candidates who volunteered and were selected to enter pilot training for manned aircraft but, due to vacancies and the prioritized need to fill RPA pilot slots, were involuntarily assigned to RPA pilot training. This group of pilot training candidates was not self-selected.

In total, 36 RPA pilot training candidates who completed UPT but involuntarily cross-trained into the RPA career field were included in this study. This group consisted of 35 (97.22%) male and 1 (2.78%) female participant with an average age of 22.63 (*SD* = 1.86) years. A total of 31 training candidates reported Caucasian (86.11%) as their ethnicity, followed by 2 Asian/Pacific Islander (5.56%) and 2 Black participants (5.56%).

3.1.3 Group 3: Manned Airframe Pilot Training Candidates. Group 3 consists of pilot training candidates who self-selected to become pilots in a manned airframe. This group completed UPT and went on to specialty training in a manned fixed or rotary wing airframe (fighter/bomber, tanker/transporter, surveillance/reconnaissance, helicopter, etc.).

In total, 7,248 manned airframe pilot training candidates were included in this study. This group consisted of 6,525 (90.02%) male and 697 (9.62%) female participants with an average age of 22.75 (*SD* = 2.73) years. A total of 6,150 participants reported Caucasian (84.85%) as their ethnicity, followed by 398 Hispanic (5.49%), 290 Asian/Pacific Islander (4.00%), 178 African (2.46%), 29 Indian (American)/Eskimo/Aleut (0.40%), and 5 Arabic participants (0.07%).

3.2 Measures

The requirements for assessing cognitive ability and developing normative data include the identification of an instrument that has high reliability and validity, as well as ease and cost of administration, and that can be administered in both group and individual settings. One instrument that suits these requirements is the MAB-II [21,22], which is based upon the structure and content of the Wechsler Adult Intelligence Scale and Wechsler's theory of intelligence [23]. The MAB-II has a testing time of 100 minutes, with 10 subscales that are each 7 minutes long. Included are five verbal ability subscales—Information, Comprehension, Arithmetic, Similarities, and Vocabulary—as well as five performance-based abilities subscales—Digit Symbol, Picture Completion, Spatial Analyses, Picture Arrangement, and Object Assembly. The subscales are weighted to calculate two intelligence quotients, verbal (VIQ) and performance (PIQ). These two IQs are used to calculate the full scale IQ (FSIQ). For the general population normative scores, the standardized scores for the MAB-II subscales have a mean of 50 (*SD* = 10), and the intelligence quotients have a mean of 100 (*SD* = 15), and have been statistically corrected for age [21]. The MAB-II manual has well-documented internal consistency reliability ranging from 0.94-0.97 on the subscales and 0.92-0.95 on the intelligence quotients and test-retest reliability ranging from 0.83-0.97 on the subscales and 0.94-0.97 on the intelligence quotients [22]. Table 1 provides descriptions of the subscales.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

Table 1. MAB-II Intelligence Quotient Subscale Descriptions

Subscales	Description
	VIQ Subscales
Information	General fund of knowledge; long-term memory
Comprehension	Social reasoning and comprehension
Arithmetic	General and numerical reasoning; problem solving
Similarities	General verbal-conceptual reasoning and problem solving
Vocabulary	Flexibility and adjustment to novelty, reasoning, abstract thought, long-term memory
	PIQ Subscales
Digital Symbol	Adaptation to new set of demands; visual learning and coding, figural memory, and speed of information processing
Picture Completion	Visual attention to detail; knowledge of common objects; perceptual and analytical skills
Spatial Score	Ability to visually and mentally rotate abstract two-dimensional images of objects in different positions; figural-domain reasoning
Picture Arrangement	Visual reasoning; ability to identify a meaningful sequence; social intelligence; perceptual reasoning
Object Assembly	Visualization and visuo-construction skills; perceptual analytical skills needed to identify a meaningful object from left-to-right sequence

Note. Adapted from Chappelle, Tran, Thompson, Goodman, and Hyde [20].

3.3 Procedure

The three groups of pilot candidates in this study were administered the MAB-II as a routine part of medical flight screening prior to attending pilot training for either manned or unmanned airframes. At the time of testing, individuals in the manned airframe pilot training group were not yet informed they would cross-train into RPA. The historical cognitive aptitude data were downloaded into a spreadsheet for analyses.

3.4 Data Analysis

Means and standard deviations on the intelligence quotients and subscales were calculated for the three groups, shown in Table 2. Generalized linear mixed models for the three training candidate groups on the intelligence quotients and subscales were run. Generalized linear mixed models were chosen for these analyses to account for the unequal sample sizes and unequal variances among the three training candidate groups.

Bonferroni post hoc *t*-tests with an adjustment for multiple comparisons were run to identify differences among groups. A statistical significance level of $p \leq 0.05$ was established a priori for the post hoc *t*-tests. A two-tailed *t*-test was not considered clinically significant (i.e., two groups were not considered meaningfully different) unless the comparison was statistically significant at $p \leq 0.05$. Small (0.20 – 0.40), moderate (0.41 – 0.79), and large (0.80 and above) Hedges' g effect sizes were identified. Other comparisons with $p < 0.15$ were noted to take into account differences between groups that may be underrepresented because of the small sample size for Group 2 (RPA training candidates who completed UPT; n = 36). Analyses were run to

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

identify the minimum sample size required for each of these comparisons to meet the $p \leq 0.05$ a priori requirements.

Table 2. Pilot Training Candidate Group Means and Standard Deviations for MAB-II

Quotients and Subscales	Group 1 (n = 411) Mean (*SD*)	Group 2 UPT (n = 36) Mean (*SD*)	Group 3 (n = 7,248) Mean (*SD*)
Intelligence Quotients			
FSIQ	120.30 (6.42)	121.31 (5.76)	120.59 (6.52)
VIQ	117.55 (6.48)	119.61 (5.56)	118.55 (6.74)
PIQ	120.50 (8.34)	120.11 (7.33)	120.02 (8.13)
VIQ Subscales			
Information	64.97 (5.93)	66.67 (5.39)	65.66 (6.07)
Comprehension	58.59 (4.15)	60.25 (3.65)	59.10 (4.36)
Arithmetic	60.02 (7.16)	60.64 (6.66)	61.15 (7.14)
Similarities	59.66 (4.57)	60.64 (3.63)	59.94 (4.58)
Vocabulary	58.48 (6.38)	59.97 (7.00)	58.85 (6.65)
PIQ Subscales			
Digit Symbol	67.67 (6.36)	68.36 (4.38)	67.52 (6.13)
Picture Completion	59.01 (6.46)	61.00 (5.68)	58.80 (6.44)
Spatial Analyses	60.85 (6.21)	59.75 (6.34)	60.77 (6.75)
Picture Arrangement	52.34 (8.02)	51.39 (6.25)	52.24 (7.91)
Object Assembly	61.93 (5.30)	61.19 (4.58)	61.53 (5.28)

Note: General population mean (*SD*) is 100 (15) for FSIQ, VIQ, and PIQ and 50 (10) for subscales. Group 1: non-rated RPA training candidates, Group 2: RPA training candidates who completed UPT, and Group 3: manned airframe training candidates.

4.0 RESULTS

This study was designed to assess differences among groups of (1) all-volunteer, RPA-direct pilot training candidates; (2) RPA pilot training candidates who completed UPT but involuntarily cross-trained; and (3) manned airframe pilot training candidates on measures of the MAB-II intelligence quotients and subscales.

4.1 Descriptive Statistics

Table 2 shows the means and standard deviations. Figure 2 displays a visual comparison of the means, ranges, and scores at 2 SDs from the mean.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

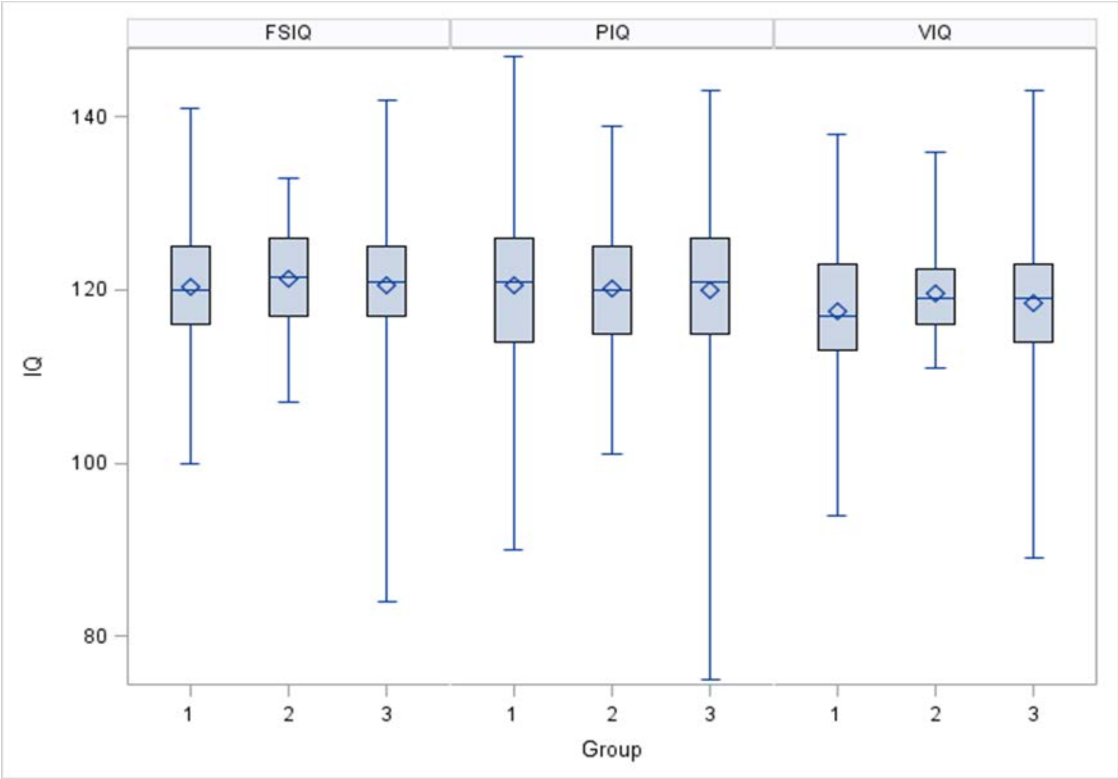

Figure 2. Comparison of IQ scores for Group 1, non-rated RPA training candidates; Group 2, RPA training candidates who completed UPT; and Group 3, manned airframe training candidates.

4.2 Assessing Between-Group Differences

Table 3 provides the general linear mixed model results for the intelligence quotients and subscales and corresponding Bonferroni post hoc *t*-tests. Significant post hoc *t*-tests with $p \leq 0.05$ are indicated in the table. Post hoc *t*-tests approaching statistical significance at $p < 0.15$ are also indicated in the table.

The VIQ ($F = 5.27$) was identified as a significant main effect. Post hoc *t*-tests identified that, on average, Group 2 (RPA pilot candidates who completed UPT) had higher VIQ scores than Group 1 (non-rated training candidates; $g = 0.32$). Group 3 also had higher VIQ scores than Group 1, but Hedges' *g* was 0.15. Main effects were identified for three verbal intelligence subscales, Information ($F = 3.33$), Comprehension ($F = 4.68$), and Arithmetic ($F = 4.82$), and one performance intelligence subscale, Picture Completion ($F = 2.85$). Post hocs identified that, on average, Group 2 had higher Comprehension scores than Group 1 ($g = 0.40$). Group 3 had higher Information, Comprehension, and Arithmetic scores than Group 1, but Hedges' *g* was 0.11, 0.12, and 0.16, respectively. Two additional significant post hoc comparisons were identified at $p < 0.15$ for Picture Completion, with Group 2 scoring higher than both Group 3 ($g = 0.34$) and Group 1 ($g = 0.31$).

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

Table 3. General Linear Mixed Models and Bonferroni Post Hoc *t*-Tests for the Three Pilot Training Groups[a]

Quotients and Subscales	Type III Tests of Mixed Effects		Group 1 (n = 411) vs. Group 2 (n = 36)				Group 1 (n = 411) vs. Group 3 (n = 7,248)				Group 2 (n = 36) vs. Group 3 (n = 7,248)			
	F test	*p*	*t*	Adj *p*	*g*	95% CI	*t*	Adj *p*	*g*	95% CI	*t*	Adj *p*	*g*	95% CI
Intelligence Quotients														
Full Scale	0.65	0.52	-0.98	0.98	-0.16	-0.33, 0.02	-0.84	1.00	-0.04	-0.09, 0.01	0.75	1.00	0.11	-0.06, 0.28
Verbal	**5.27**	**0.01**	**-2.10**	**0.11**[b]	**-0.32**	**-0.49, -0.15**	**-3.01**	**0.01**	**-0.15**	**-0.20, -0.10**	1.14	0.76	0.16	-0.01, 0.33
Performance	0.66	0.52	0.30	1.00	0.05	-0.13, 0.22	1.15	0.75	0.06	0.01, 0.11	0.08	1.00	0.01	-0.16, 0.18
VIQ Subscales														
Information	**3.33**	**0.04**	-1.80	0.21	-0.29	-0.46, -0.11	**-2.30**	**0.06**	**-0.11**	**-0.16, -0.06**	1.12	0.79	0.17	0.00, 0.33
Comprehension	**4.68**	**0.01**	**-2.57**	**0.03**	**-0.40**	**-0.58, -0.23**	**-2.37**	**0.05**	**-0.12**	**-0.17, -0.07**	1.88	0.18	0.26	0.10, 0.43
Arithmetic	**4.82**	**0.01**	-0.52	1.00	-0.09	-0.26, 0.09	**-3.08**	**0.01**	**-0.16**	**-0.21, -0.11**	-0.46	1.00	-0.07	-0.24, 0.10
Similarities	1.36	0.26	-1.49	0.41	-0.22	-0.39, -0.04	-1.16	0.73	-0.06	-0.11, -0.01	1.14	0.76	0.15	-0.02, 0.32
Vocabulary	1.16	0.31	-1.24	0.65	-0.23	-0.41, -0.06	-1.17	0.73	-0.06	-0.11, -0.01	0.96	1.00	0.17	0.00, 0.34
PIQ Subscales														
Digit Symbol	0.78	0.46	-0.86	1.00	-0.11	-0.29, 0.06	0.51	1.00	0.02	-0.03, 0.07	1.15	0.75	0.14	-0.03, 0.30
Picture Completion	2.85	0.06	**-1.99**	**0.14**[c]	**-0.31**	**-0.48, -0.14**	0.63	1.00	0.03	-0.02, 0.08	**2.31**	**0.06**[d]	**0.34**	0.17, 0.51
Spatial Analyses	0.52	0.59	1.02	0.92	0.18	0.00, 0.35	0.33	1.00	0.01	-0.04, 0.06	-0.96	1.00	-0.15	-0.32, 0.02
Picture Arrangement	0.36	0.70	0.84	1.00	0.12	-0.05, 0.29	0.22	1.00	0.01	-0.04, 0.06	-0.81	1.00	-0.11	-0.27, 0.06
Object Assembly	1.23	0.29	0.81	1.00	0.14	-0.03, 0.31	1.52	0.39	0.08	0.03, 0.13	-0.38	1.00	-0.06	-0.23, 0.10

Note. Columns 1 and 2 are univariate mixed model effects ($p \leq 05$). Bonferroni post hoc *p*-values are adjusted for multiple comparisons ($p \leq 05$). Statistical analyses with $p \leq .05$ are in darker shaded boxes. Statistical analyses with *p* between .06 and .15 (approaching statistical significance) are shown in lighter shaded boxes. Hedges' *g* effect size ranges: small (0.20 - 0.40), moderate (0.41 - 0.79), and large (0.80 and above).
CI = confidence interval.
[a]Group 1: non-rated RPA training candidates; Group 2: RPA training candidates who completed UPT; and Group 3: manned airframe training candidates. Minimum sample size required for each group for mean comparison to approach statistical significance at $p \leq 05$:
[b]n = 136.
[c]n = 148.
[d]n = 120.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

5.0 DISCUSSION

This study compared the MAB-II scores for three distinct groups of military aviation trainees: those who self-selected to go straight into RPA pilot training (Group 1), those who self-selected to attend UPT and were assigned to an RPA platform upon graduation from UPT (Group 2), and those who self-selected to attend UPT and were assigned to a manned aircraft upon graduation (Group 3). The MAB-II was administered as part of routine medical screening for all three groups prior to entering their initial training (UPT or RPA).

Overall, all three groups scored in the superior range (at over above the 91st percentile) for FSIQ and PIQ and in the high average to superior range (between the 87th and 90th percentiles) for VIQ when compared with peers of similar age in the general population (Table 2). This is consistent with prior research on USAF manned and unmanned pilot populations [4,20,24,29]. Furthermore, the results of the study revealed the MAB-II test scores for these three pilot groups had significantly less variance when compared with age-corrected normative general population scores (standard deviation for the three quotients ranged from 37.0-55.6% of the normative value; see Table 2). Simply put, the scores among these three pilot training candidate groups were more homogenous with a much smaller range than the distribution and range of scores for their peers in the general population. Again, this is consistent with prior research on aviator populations [24]. It is well documented that aviation career fields require high levels of cognitive ability, and RPA pilot candidates (Groups 1 and 2) appear to be no exception, regardless of the source of the trainees.

<u>Group 1 vs. Group 2</u>: No differences were detected between FSIQ and PIQ scores between non-rated RPA pilot training candidates (Group 1) and RPA pilot training candidates who completed UPT (Group 2). Both groups, on average, scored in the "superior" range of intellectual functioning for both FSIQ and PIQ, as compared to their age-normed peers in the general population.

For VIQ, both groups scored in the "high average" range as compared to the general population. This is consistent with prior research on both manned and unmanned pilot groups [20,29]. However, a small, but statistically significant, difference in VIQ (primarily accounted for by the Comprehension subscale) was found between these two groups. RPA training candidates who completed UPT scored slightly higher than non-rated RPA training candidates (Table 2). Although the effect size for the VIQ difference approaches the moderate range (g=0.32; 95% CI = 0.15 – 0.49), this finding represents small differences in VIQ at the high end of functioning. Small differences between groups that are in the high range of functioning (i.e., concerns regarding possible ceiling effect) and whose distributions have significant overlap are difficult to interpret and may not represent practical differences in verbal intelligence between these two groups. Future research should focus on the predictive ability of VIQ scores for RPA-relevant training and performance outcomes to determine if this small difference in VIQ has operational significance.

Interpreting the difference found between Group 1 and Group 2 on the Comprehension subscale presents similar concerns, especially in light of the belief that those RPA training candidates who volunteer directly for RPA training (Group 1) may be "video-gamers" whose general social reasoning and social disposition may not be as sophisticated. The Comprehension subscale has been thought to measure social reasoning and comprehension, the "social knowledge base" of societal conventions and norms [25]. However, prior research attempting to operationalize this belief has failed to consistently find associations between the Comprehension

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

subscale and other validated measures of social functioning, beyond the well-established association between general intelligence and social functioning [26]. Additionally, the application of this previous research to the current population is limited by many factors. The research to date has focused almost exclusively on populations known to have social deficits (e.g., schizophrenia, attention deficit hyperactivity disorder, and autism spectrum disorders) and not on very high functioning populations such as aviators. Moreover, prior research has indicated the importance of relative deficits in Comprehension scores as compared to a subject's general intelligence in the relationship between the Comprehension subtest and social functioning [27,28]. Lastly, the Wechsler tests of intellectual ability have been used almost exclusively in these previous studies. While the Wechsler tests and the MAB-II Comprehension subscales are highly correlated (MAB/Wechsler Adult Intelligence Scale-R Comprehension subscale r = .73 [22]), no studies could be found regarding the MAB-II Comprehension subscale and an established measure of social knowledge and functioning. The deficit-focused approach used by previous authors to describe the relationship among the Comprehension subscale, general intelligence, and social functioning is not generalizable to the current data that clearly involve two groups who scored, on average, in the high average to superior range of functioning on the Comprehension subscale as well as on general intelligence. Lastly, in the context of high general intelligence, and in the absence of relative deficits on this subscale, the between-group difference on the Comprehension subscale is highly unlikely to represent a functional difference in social knowledge or skills between these two groups.

The last statistically significant difference between Groups 1 and 2 was found on the Performance subscale of Picture Completion. This subtest requires subjects to identify what is missing in a series of sketch-type pictures. Successful performance of this task requires knowledge of a variety of common objects and the perceptual and analytic skills to interpret the picture and distinguish important details from unessential omissions. On average, both of these groups scored between 0.9 and 1.1 SDs above the mean for the age-corrected normative population on this subtest (equivalent to the 82nd percentile and the 86th percentile). Again, considering that both groups scored in the high average range and well above the general population mean of 50, it is unlikely that this difference would lead to notable differences in perceptual and analytic abilities during training of flight.

Overall, trainees sent to become RPA pilots after UPT (Group 2) do not appear to be different, intellectually, than those who self-select to go directly to RPA pilot training (Group 1). Both groups scored in the superior range for FSIQ and PIQ and in the high average range for VIQ, indicating that regardless of whether they are selected from UPT graduates or from the volunteer pool of recently commissioned officers, RPA pilot trained candidates performed very well on this general measure of intelligence.

Group 1 vs. Group 3: Although there were some statistically significant differences between Groups 1 and 3 (non-rated RPA pilot training candidates and manned aircraft pilot training candidates, respectively) on the VIQ subscales of Information, Comprehension, and Arithmetic (and as a result, the VIQ overall), the effect sizes (g = 0.10-0.21; see Table 3) are so small as to bring into question any practical significance for these findings. Both groups, on average, scored in the high average to superior range of intellectual functioning for all three intelligence quotients. These data suggest that non-rated RPA pilot trainees are very similar intellectually to those trainees who attend UPT and manned pilot aircraft training. Additionally, the results are incompatible with the existing stereotype that those who desire to pilot an RPA to

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

the exclusion of flying manned aircraft are less capable (intellectually) than those who are chosen to pilot a manned aircraft.

Group 2 vs Group 3: A small but statistically significant difference on the Picture Completion subscale was found between RPA pilot trainees who completed UPT (Group 2) and manned airframe pilot trainees (Group 3), with Group 2 scoring higher than Group 3 (mean = 61.0 [*SD* = 5.68; 86th percentile] and mean = 58.8 [*SD* = 6.44; 82nd percentile], respectively). The Picture Completion subscale measures visual attention to detail and perceptual and analytical skills including the ability to differentiate essential from non-essential details [21]. Similar to the previous differences discussed, this finding must also be interpreted with caution. Although the effect size approaches moderate ($g = 0.34$; 95% CI = 0.17, 0.51), these are two groups whose distributions significantly overlap and whose means, for both FSIQ and for this subtest, fall at the high average to superior range of functioning. From a practical sense, this difference likely does not represent a meaningful difference in visual attention to detail or perceptual and analytic skills between these two groups. However, future research into the unique skills and abilities required of the RPA career field should further investigate this difference in the context of other factors potentially predictive of training outcomes and performance (e.g., motivation, personality, prior skills, and knowledge, etc.).

Overall, the results of this study suggest that these three pilot training groups are intellectually more similar than they are different. There were no indications that those in training to become RPA pilots, whether self-selected (Group 1) or assigned after UPT (Group 2), are intellectually different from each other or manned pilot trainees (Group 3). All three groups performed as would be expected based upon prior research on aviation communities: in the superior range of functioning for FSIQ and PIQ and in the high average range for VIQ.

5.1 Limitations and Areas for Future Research

The current study describes and compares three groups of trainees who entered the training pipeline for manned or unmanned aircraft within the specified window in an effort to identify potential selection biases or differences that might exist between the manned and unmanned training pipelines. This study did not identify or compare those that passed training for their assigned platform with those who did not. Future research should seek to replicate the current findings and determine whether the subtle differences found in this study have predictive ability for training and performance outcomes or for suitability for specific aviation platforms. As a number of prior studies have shown, understanding and predicting aviator performance is a complex mix of ability, personality, and motivation [7]. Predicting performance may require unique algorithms for type of aircraft flown – be they manned or unmanned [24], and subtle differences on subscales may differentiate between those who succeed and those who fail in training [20]. Although overall these groups appear very similar intellectually, the small but statistically significant differences seen in this study may have predictive utility when considered in the context of an aviator's motivation, personality, and physical ability. King et al. [7] found that in a similar, restricted range, high functioning study of military pilots, mean FSIQ MAB-II scores between those who graduated training and those who were eliminated from training differed by less than two IQ points. Future research will tell if the differences identified in this research will have performance implications.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

The sample size for Group 2 (RPA pilot trainees who completed UPT) is a primary limitation to the current study. As low statistical power undermines both the ability to detect a true effect and the reliability of effects found, our low Group 2 sample size encourages future replication of these results. However, the fact that the MAB-II scores for Group 2 are consistent with prior research on aviators and are so similar to the other two groups in this study, the current results likely represent an accurate picture of the intellectual ability of these three groups.

Lastly, the MAB-II scores in the current study were significantly restricted in range, which is a known consequence of the necessary selection bias for these very high functioning personnel. In the future, it would be beneficial to identify a measure of cognitive functioning that provided better discriminability in the high ability range.

5.2 Conclusion

The current study represents an important step in eliminating the negative stereotypes that have evolved regarding RPA trainees. RPA trainees, whether they are acquired from UPT graduates or from newly commissioned officers in the USAF, performed exceptionally well on a standardized test of intellectual ability. Additionally, they performed very similarly to UPT trainees who eventually went on to fly manned aircraft, suggesting that based on intellectual ability, none of these groups of aviation trainees are superior to the other. Future research should examine other cognitive tests, as well as indicators of personality and motivation, to determine whether any significant differences exist among these groups and, if found, if those differences reflect operationally relevant skills and abilities.

6.0 REFERENCES

1. U.S. Air Force. MQ-1B Predator [Fact sheet]. 2010 Jul 20. [Accessed 27 Mar 2015]. Available from http://www.af.mil/AboutUs/FactSheets/Display/tabid/224/Article/104469/mq-1b-predator.aspx.
2. U.S. Air Force. MQ-9 Reaper [Fact sheet]. 2010 Aug 18. [Accessed 27 Mar 2015]. Available from http://www.af.mil/AboutUs/FactSheets/Display/tabid/224/Article/104470/mq-9-reaper.aspx.
3. Paullin C, Katz L, Bruskiewicz KT, Houston J, Damos D. Review of aviator selection. Arlington (VA): U.S. Army Research Institute for the Behavioral and Social Sciences; 2006. Technical Report 1183.
4. Callister JD, King RE, Retzlaff PD. Cognitive assessment of USAF pilot training candidates. Aviat Space Environ Med. 1996; 67(12):1124–1129.
5. Ree MJ, Carretta TR. Central role of g in military pilot selection. Int J Aviat Psychol. 1996; 6(2):111-123.
6. Chappelle W, Ree MJ, Barto EL, Teachout MS, Thompson WT. Joint use of the MAB-II and MicroCog for improvements in the clinical and neuropsychological screening and aeromedical waiver process of rated USAF pilots. Brooks City-Base (TX): U.S. Air Force School of Aerospace Medicine; 2010. Technical Report No. AFRL-SA-BR-TR-2010-0002.
7. King RE, Carretta TR, Retzlaff P, Barto E, Ree MJ, Teachout MS. Standard cognitive psychological tests predict military pilot training outcomes. Aviat Psychol Appl Hum Factors 2013; 3(1):28-38.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

8. Thompson WT, Orme DR, Zazeckis TM. Neuropsychological evaluation of aviators: need for aviator-specific norms? Brooks City-Base (TX): U.S. Air Force School of Aerospace Medicine; 2004. Technical Report SAM-FE-BR-TR-2004-0001.
9. Schmidt FL, Hunter JE. The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. Psychol Bull. 1998; 124(2):262-274.
10. Schmidt FL, Hunter J. General mental ability in the world of work: occupational attainment and job performance. J Pers Soc Psychol. 2004; 86(1):162-173.
11. Carretta TR, Ree MJ. U.S. Air Force pilot selection tests: what is measured and what is predictive. Aviat Space Environ Med. 1996; 67(3):279-283.
12. Martinussen M. Psychological measures as predictors of pilot performance: a meta-analysis. Int J Aviat Psychol. 1996; 6(1):1-20.
13. Pavlas D, Burke CS, Fiore SM, Salas E, Jensen R, Fu D. Enhancing unmanned aerial system training: a taxonomy of knowledge, skills, attitudes, and methods. Human Factors and Ergonomics in Society Annual Meeting Proceedings. 2009; 53(26):1903-1907.
14. Biggerstaff S, Blower DJ, Portman CA, Chapman AD. The development and initial validation of the unmanned aerial vehicle (UAV) external pilot selection system. Pensacola (FL): Naval Aerospace Medical Research Laboratory, 1998. Report No. NAMRL-1398.
15. Kay G, Dolgin D, Wasel B, Langelier M, Hoffman C. Identification of the cognitive, psychomotor, and psychosocial skill demands of uninhabited combat aerial vehicle (UCAV) operators. Patuxent River (MD): Naval Air Warfare Center, Aircraft Division; 1999.
16. Nagy JE, Muse K, Eaton G, Phillips A. U.S. Air Force unmanned aircraft systems performance analyses: Global Hawk pilot and sensor operator front end analysis (FEA) report [Access controlled]. Wright-Patterson AFB (OH): Survivability/Vulnerability Information Analysis Center; 2007. SURVIAC-TR-10-041.
17. Nagy JE, Kalita SW, Eaton G. U.S. Air Force unmanned aircraft systems performance analyses: Predator pilot front end analysis (FEA) report [Access controlled]. Brooks City-Base (TX): 311th Human Systems Wing, Performance Enhancement Directorate; 2006. Technical Report SURVIAC-TR-06-203.
18. Bailey M. Predator pilot and sensor operator selection test batteries. Cranwell Royal Air Force Base (UK): Royal Air Force; 2009. [Available by request only].
19. Chappelle W, McDonald K, McMillan K. Important and critical psychological attributes of USAF MQ-1 Predator and MQ-9 Reaper pilots according to subject matter experts. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine; 2011. Technical Report AFRL-SA-WP-TR-2011-0002.
20. Chappelle W, Tran NVT, Thompson W, Goodman T, Hyde K. Intelligence and neuropsychological aptitude testing of U.S. Air Force MQ-1 Predator pilot training candidates. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine; 2012. Technical Report AFRL-SA-WP-TR-2013-0003.
21. Jackson DN. Multidimensional Aptitude Battery-II: manual. Port Huron (MI): Sigma Assessment Systems, Inc.; 2003.
22. Jackson DN. Multidimensional Aptitude Battery-II: technical manual. Port Huron (MI): Sigma Assessment Systems, Inc.; 2013.
23. Wechsler D. Wechsler Adult Intelligence Scale-3rd Edition (WAIS-3®). San Antonio (TX): Harcourt Assessment; 1997.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

24. Boyd JE, Patterson JC, Thompson BT. Psychological test profiles of USAF pilots before training vs. type aircraft flown. Aviat Space Environ Med. 2005; 76(5):463-468.
25. Sipps GJ, Berry GW, Lynch EM. WAIS-R and social intelligence: a test of established assumptions that uses the CPI. J Clin Psychol. 1987; 43(5):499-504.
26. Campbell JM, McCord DM. Measuring social competence with the Wechsler Picture Arrangement and Comprehension subtests. Assessment. 1999; 6(3):215-224.
27. Allen DN, Strauss GP, Donohue B, van Kammen DP. Factor analytic support for social cognition as a separable cognitive domain in schizophrenia. Schizophr Res. 2007; 93(1-3):325-333.
28. Goldstein G, Beers SR, Siegel DJ, Minshew NJ. A comparison of WAIS-R profiles in adults with high-functioning autism or differing subtypes of learning disability. Appl Neuropsychol. 2001; 8(3):148-154.
29. King RE, Barto E, Ree MJ, Teachout MS, Retzlaff P. Compilation of pilot cognitive ability norms. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine; 2011. Technical Report AFRL-SA-WP-TR-2012-0001.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015

LIST OF ABBREVIATIONS AND ACRONYMS

CI confidence interval

FSIQ full scale intelligence quotient

MAB-II Multidimensional Aptitude Battery-Second Edition

PIQ performance intelligence quotient

RPA remotely piloted aircraft

SD standard deviation

UPT undergraduate pilot training

USAF U.S. Air Force

VIQ verbal intelligence quotient

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2255, 5 May 2015