SHORT COMMUNICATION

PREDICTIVE VALIDITY OF THE PSYCHOLOGICAL ASSESSMENT OF ADOLESCENTS PLANNING TO START AVIATION TRAINING. PRELIMINARY STUDY

Piotr ZIELIŃSKI

Department of Aviation Psychology, Military Institute of Aviation Medicine, Warsaw, Poland

Source of support: Own sources

Author's address: P. Zieliński, Military Institute of Aviation Medicine, Krasińskiego Street 54/56, 01-755 Warsaw, e-mail: pzielins@wiml.waw.pl

Introduction: Cognitive and psychomotor skills as well as selected personality traits are important factors determining success in aviation training. Many people decide to take the aviation career at an early age, so the question arises whether the psychological features visible in developmental age can be a basis for predicting success in adulthood.

Methods: The analysis took into account the results of psychological tests of 97 people (15-16 years old) who were enrolled in the Aviation High School, then after graduation again tested as candidates for aviation studies. Measurements of the stability of the psychological variables were calculated, as well as the factor structure of the results. Then, the predictive value of factors distinguished in this analysis was determined.

Results: Despite the visible developmental changes in cognitive functions (higher results in the second assessment), most of the indicators are characterized by a satisfactory stability. In the structure of the results, six relatively independent factors were distinguished, one of which, defined as general intellectual ability, turned out to be the best predictor of a positive qualification in later assessment.

Discussion and Conclusions: The obtained results allow us to conclude that the psychological tests and indicators used in the study of adolescents seem to be adequate and guarantee high stability of measurements, even over a period of several years. As predicted, the greatest predictive power has the general intellectual ability, which is the basic property that allows the acquisition and consolidation of other skills.

Keywords: cognitive performance, developmental changes, factor analysis, measurement stability, psychomotor performance

INTRODUCTION

An important part of the medical examination for admission to aviation training, both in military and civil (professional) aviation, is psychological testing. The areas diagnosed are those considered key in the pilot's profession, i.e. both basic mental competencies (reasoning, memory, etc.) and specific operational skills (psychomotor skills, multitasking, etc.) and competencies based on personality traits. Specifically, it should be stated that the following traits and functions relevant to the pilot profession are distinguished within these areas in aviation psychology [14,15]:

- **reasoning** – a cognitive process that refers to finding and applying general principles and logical relationships in various task situations and problem solving;
- **memory functions** – the ability to remember and retrieve visual and/or auditory information from memory, even in the presence of distractions;
- **attention processes**, including:
    - **focus** – the ability to direct attention to ensure a stable level of task completion;
    - **vigilance** – the ability to maintain a state of readiness for long periods of time in order to respond efficiently to infrequent, irregular occurrences;
    - **divisibility** – the ability to effectively direct attention simultaneously to different tasks;
    - **selectivity** – the ability to selectively direct attention to a chosen type of information despite the presence of distractions;
- **perception** – the ability to quickly perceive and interpret sensory information;
- **spatial abilities** – the ability to correctly construct mental spatial images and to correctly perceive the relationships between objects in space;
- **psychomotor functions** – the speed of response to stimuli and the ability to coordinate movements in response to dynamic sensory information;
- **personality factors related to functioning at work** – the tendency to arouse and

maintain motivation to achieve the desired goal while maintaining a positive attitude towards the job (e.g. need for achievement, conscientiousness, openness);

- **social competence** – the ability to form and maintain good relationships with other people, (including extraversion, dominance, empathy, aggressiveness);
- **emotional stability** – the tendency to respond to difficult and/or threatening situations in a manner that is adequate and provides emotional control;

Numerous studies over the past several decades (for a historical outline of earlier studies, see, e.g., Griffin and Koonce, 1996 [9]) demonstrate the significant predictive power of many of these areas in predicting success in flight training [8,18,29], and meta-analyses highlight the particular importance of factors such as spatial ability or perceptual speed, especially relative to verbal ability or personality traits, the measurement of which has less predictive accuracy in the aviation domain [2,13]. A characteristic of all of these studies, however, is that the psychological measurement is treated pointwise, as a single test prior to flight training. This implies an assumption of constancy over time of the measured traits, while it is otherwise known that many cognitive functions are still subject to development in early adulthood [11], and their level can change significantly even in adulthood as a result of specific experiences or training [27].

The Military Institute of Aviation Medicine (WIML) has a unique opportunity to provide psychological testing at various stages of an aviation career. The first tests are being conducted on teenagers, candidates for the General Aviation High School (OLL) in Dęblin. Subsequently, many of these individuals (both OLL graduates and those who eventually graduated from another high school) are reexamined as candidates for the Polish Air Force University (LAW). The results collected in this way provide an important knowledge base regarding the formation of special abilities and the possibility of predicting the level of professionally important traits based on developmental outcomes by assessing the constancy over time of these measurements.

The analysis presented in this article is based on data collected on first-year candidates for OLL who were tested at WIML. It has a dual purpose. First, it will serve to determine the stability of the test indicators used within those tests that are repeated in the same or related form during candidate testing in adolescents and adults. The second, more important goal is to isolate, based on the test indicators collected, the underlying factors that form coherent diagnostic domains and then evaluate the predictive power of these factors in predicting performance during LAW candidate testing.

Such formulation of the research problems will allow us to answer the question of which areas most reliably allow – even taking into account the developmental changes visible in their level – to predict a positive assessment of professional suitability in people wishing to begin a professional career in aviation.

METHODS

Test subjects

The study using *ex post facto* model used data obtained during initial testing of OLL candidates from 102 test subjects (15 females) aged 15-16 years (M=15.29, SD=0.48) who then participated in initial testing of LAW candidates (mean age during testing for LAW was 18.2, SD=0.42). After eliminating the missing data, 97 individuals remained in this group, 76 of whom were OLL graduates (individuals who, after psychological testing and qualification, gained admission to OLL and, after graduation from that school, were candidates for LAW) and 21 were candidates for LAW after graduating from another high school (all individuals in the subgroup of 21 also passed psychological testing for OLL; there is no information in the data analyzed as to the reason why they ultimately did not attend OLL and ended up at another school).

Indicators

Test indicators from both questionnaire studies (self-description on personality traits) and performance tests on various areas of cognitive and psychomotor performance were used in the analysis. Specifically, these were:

1) The results of the *NEO-FFI* test [33], used in the tests of OLL candidates, and its longer version *NEO-PI-R* [28], used in the tests of LAW candidates, which means five dimensions included in Costa and McCrae's personality theory, namely Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness [5].
2) The results of the *IVE* test [16], used to determine personality dimensions such as Impulsivity, Venturesomeness, and Empathy.
3) The results of the *APIS-P(R)* test [20], a battery used to measure the profile of crystallized intelligence in adolescents. This battery includes subtests measuring social, verbal, abstract-logical, and visual-spatial skills.
4) The results of the Raven's Advanced Progressive Matrices (in Polish: Test Matryc

Ravena w wersji dla Zaawansowanych) [17], used to measure fluid intelligence in high-functioning individuals.

5) The results of *PST* test from the Vienna Test System [10], indicating the level of spatial abilities.
6) The results of *CORSI* test from the Vienna Test System [26], which determines the operating memory capacity.
7) The results of *COG* test from the Vienna Test System [32], i.e., speed and correctness of responses; the *COG* test is used to measure perceptual speed and efficiency of attention processes.
8) The results of *DT* test from the Vienna Test System [21], used to assess speed and motor control.
9) The results of *B19* test from the Vienna Test System [4], used to measure psychomotor coordination.
10) The results of *SMK* test from the Vienna Test System [3], used to measure psychomotor coordination. Due to the level of difficulty, the *B19* test is used for OLL candidates and the *SMK* test is used for LAW candidates.

Statistical analysis

The first step in the statistical analysis was to analyze the stability of test results over a three-year interval (time between initial testing of OLL candidates and LAW candidates). Due to the different characteristics of the distributions of the individual test indicators, the analysis was performed using non-parametric statistics, including Spearman's *rho* correlation coefficient. It allows the detection of monotonic relationships (steadily increasing or steadily decreasing), without assuming the linearity of such a relationship (which linearity, in the case of results that determine developmental age changes, would probably be an unrealistic expectation).

In the next step, an exploratory factor analysis was conducted to isolate the underlying latent variables behind the test results obtained in the tests of OLL candidates. Once these factors were identified, a regression model was used to determine which factors were most significantly predictive of positive psychological test scores during testing of LAW candidates.

All statistical analyses were performed in *R* statistical environment version 3.4.0 [22] using *psych* package version 1.7.5 [24].

RESULTS

Stability analysis

For the analysis of so-called absolute stability, which is the assessment of the amount of change in results obtained by using the same psychological test, with an interval of approximately three years between each test, *COG s1*, *DT s5*, and *CORSI s1* tests were included. The main test indicators were analyzed. Both the difference and correlation between the two measurements were assessed. Because the distributions varied in terms of normality, non-parametric statistics (median, interquartile deviation, Spearman's rank correlation coefficient, and Wilcoxon test) were used. When there are significant differences between two measurements, the correlation coefficient should obviously not be taken as a measure of the tool reliability. The results are presented in Table 1.

Table 1. Absolute stability of selected tests used for testing OLL candidates (n=97).

Indicator	**MED 1**	**IQR 1**	**MED 2**	**IQR 2**	**Correlation (*rs*)**	**Correlation (*p*)**	**Difference (Wilcoxon *p*)**	**Effect size (Wilcoxon *r*)**
CORSI (result)	6	1	6	0.5	0.34	<0.001	<0.001	0.33
CORSI (accurate)	11	1.5	13	1.5	0.41	<0.001	<0.001	0.36
COG (time of rejected)	2.09	0.25	1.72	0.2	0.61	<0.001	<0.001	0.52
COG (time of accepted)	1.94	0.21	1.68	0.21	0.57	<0.001	<0.001	0.5
COG (total rejected)	113	3	115	2	0.52	<0.001	<0.001	0.3
COG (total accepted)	74	3	75	2	0.2	0.05	0.11	0.14
DT (int1 responses over time)	170	5.5	178	1.5	0.37	<0.001	<0.001	0.51
DT (int2 responses over time)	114	24.5	165	7.5	0.63	<0.001	<0.001	0.59
DT (int3 responses over time)	150	15.5	172	3.25	0.53	<0.001	<0.001	0.58
DT (int1 number of errors)	7	3	4	2	0.13	0.22	<0.001	0.3
DT (int2 number of errors)	17	6	8	3.5	0.15	0.15	<0.001	0.45
DT (int3 number of errors)	13	5	7	3	0.27	0.01	<0.001	0.46
DT (difference int1 – int2)	54	24	11	6.5	0.56	<0.001	<0.001	0.58
DT (difference int1 – int3)	20	11	5	3.5	0.37	<0.001	<0.001	0.54

Note. MED1, MED2 – median scores in the first (OLL) and second (LAW) tests; IQR1, IQR2 – interquartile deviation in the first (OLL) and second (LAW) tests.

The results presented in the table above show a moderate but satisfactory level of stability of the tests used (from 0.34 to 0.63), only the error rates in the *DT* test and the number of correctly accepted stimuli in the *COG* test obtain lower values of the correlation coefficient. In the first case, this result is related to the generally low accuracy of error rates in the *DT* test and the very high susceptibility to situational factors, while in the second case the low correlation is due to the very high consistency and low variance of the results in the analyzed group (small range of the indicator).

Next, we analyzed the compliance of the results of related tests, whose different versions (e.g. due to their adaptation to the age of the test subjects) are used during testing of candidates for OLL and LAW. Thus, we compared the *NEO-FFI* test indicators (testing for OLL) with the main factor scores of the *NEO-PI-R* test (testing for LAW), and the total score and individual subtest scores of the *APIS-P(R)* test (testing for OLL) with results of the *TMZ* and *PST* tests (testing for LAW), and the total indicators of the *B19* test (testing for OLL) with the main indicators of the *SMK s2* test (testing for LAW). Because these tests have different structure and the raw scores are calculated differently, it is pointless to calculate the differences between their scores. Therefore, the analysis was limited to assessing Spearman's *rho* monotonic correlation. The results are presented in Table 2.

Table 2. Relative stability of selected tests used for testing OLL candidates (n=97).

OLL testing \ LAW testing	Neuroticism (NEO-PI-R)	Extraversion (NEO-PI-R)	Openness (NEO-PI-R)	Agreeableness (NEO-PI-R)	Conscientiousness (NEO-PI-R)	TMZ	PST	SMK
Neuroticism (NEO-FFI)	0.21 *							
Extraversion (NEO-FFI)		0.4 *						
Openness (NEO-FFI)			0.34 *					
Agreeableness (NEO-FFI)				0.43 *				
Conscientiousness (NEO-FFI)					0.35 *			
Impulsivity (IVE)	0.28 *	0.03						
Empathy (IVE)	-0.03	0.21 *						
Venturesomeness (IVE)	-0.01	0.12						
APIS-P(R) Overall score						0.42 *	0.21 *	
APIS-P(R) Test 1						0.19	0.2 *	
APIS-P(R) Test 2						0.3 *	0.13	
APIS-P(R) Test 3						0.16	0.09	
APIS-P(R) Test 4						0.17	-0.03	
APIS-P(R) Test 5						0.42 *	0.33 *	
APIS-P(R) Test 6						0.06	0.04	
APIS-P(R) Test 7						0.29 *	0.07	
APIS-P(R) Test 8						0.18	0.16	
B19 error time								-0.31 *
B19 number of errors								-0.3 *

Note. The Table includes Spearman's rho rank correlation coefficients; * $p < .05$.

Within the personality traits, the analysis shows moderate predictive power of the *NEO-FFI* test in relation to the *NEO-PI-R* test results. The extraversion and agreeableness scales were most strongly correlated, while the neuroticism scale showed the weakest correlation. Motor tests *(B19* and *SMK*), despite their different structure, show a moderate positive relationship. Tests related to intelligence and spatial ability show satisfactory convergent and discriminant validity. The results of the *TMZ* test (measuring fluid intelligence) correlate in an acceptable manner ($rs>0.4$) with the overall score of the *APIS-P(R)* test (crystallized intelligence), and of the tests comprising the profile of skills – with the test for measuring abstract-logical abilities, based on abstract material, which is the most related area to the so-called fluid intelligence. In contrast, as expected, *TMZ* test scores correlate lowest with culturally loaded areas of intelligence, such as vocabulary and social tests, highlighting the accuracy of the diagnosis of the distinguished traits. A somewhat surprising result is the weak relationship of the measure of fluid intelligence with abstract-logical abilities based on concrete material, which may suggest that Test 4 in the *APIS-P(R)*, contrary to its assignment to abstract-logical abilities, measures the level of school knowledge rather than the actual ability to classify and find logical connections. Contrary to expectations, the *PST* from the battery of LAW candidates used to measure spatial ability (the accuracy of which is proven in the examination of pilots) does not show convergence with visual-spatial test results from the *APIS-P(R)* battery.

Exploratory factor analysis

The next step, after assessing the stability of the results, was *exploratory factor analysis* (EFA), which aimed to extract, within the analyzed indicators from the testing for OLL (cf. Table 1 and 2), the main psychological constructs that can then serve as predictors of positive outcome in LAW candidate testing.

First, the optimal number of factors comprising the various indicators included in the analysis was determined. Algorithms for estimating the optimal number of factors contained in the *VSS* and *fa.parallel* procedures from the *psych* package [24] were used, including the *Very Simple Structure* algorithm [25], the *MAP* algorithm [31], and parallel analysis [12] with *varimax* rotation estimated by the maximum likelihood method.

The parallel analysis suggests a six-factor solution; the scree plot narrows this suggestion to five factors. In contrast, the *MAP* criterion suggests 4 factors, although such a model has an unsatisfactory level of fit ($RMSEA > 1$), and the *VSS* criterion suggests as many as eight independent factors. Satisfactory level of *RMSEA* is achieved by 6-factor and more complex models, hence it was finally decided to estimate a six-factor model. This solution, after applying Kaiser normalization, is presented below (Table 3):

Table 3. Factor loadings for the set of indicators obtained in the tests for OLL (n=97).

Indicator	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**	**Unique Value**
CORSI memory span	0.98	0.04	-0.11	0.09	0.13	0.06	0.005
CORSI Number of accurate	0.84	0.16	-0.23	0.09	0.06	-0.02	0.197
COG Time of rejected	-0.27	-0.2	0.91	-0.21	-0.08	-0.1	0.005
COG Time of accepted	-0.29	-0.16	0.79	-0.25	-0.13	-0.03	0.188
COG Total rejected	0.11	0.07	0.02	-0.01	-0.3	-0.05	0.978
COG Total accepted	0.07	0.14	0.49	-0.05	-0.05	0.08	0.722
DT int1 Responses over time	-0.08	-0.08	-0.06	0.28	-0.09	-0.12	0.881
DT int2 Responses over time	0.2	0.03	-0.16	0.93	-0.18	0.02	0.032
DT int3 Responses over time	0.11	0.08	-0.17	0.82	-0.17	0.02	0.247
DT int1 Number of errors	0.14	-0.11	-0.18	0.04	0.65	-0.03	0.504
DT int2 Number of errors	0.01	0.00	-0.03	-0.23	0.76	-0.03	0.368
DT int3 Number of errors	0.00	-0.05	-0.01	-0.24	0.91	-0.02	0.118
NEO-FFI Neuroticism	-0.13	-0.01	0.07	0.09	-0.03	-0.32	0.867
NEO-FFI Extraversion	-0.13	0.16	0.05	-0.11	0.08	0.18	0.906
NEO-FFI Openness	-0.11	0.19	0.08	0.05	0.17	0.37	0.774
NEO-FFI Agreeableness	-0.03	-0.11	0.11	-0.01	-0.06	0.86	0.236
NEO-FFI Conscientiousness	0.02	-0.18	-0.08	0.06	-0.04	0.46	0.746
APIS-P(R) Overall score	0.18	0.98	-0.02	0.01	-0.06	-0.01	0.005
APIS-P(R) Test 1	-0.02	0.54	0.04	-0.01	0.00	0.29	0.625
APIS-P(R) Test 2	0.23	0.53	0.08	-0.07	0.12	0.12	0.624
APIS-P(R) Test 3	-0.01	0.64	-0.04	0.01	-0.12	0.04	0.575
APIS-P(R) Test 4	0.18	0.52	-0.05	-0.18	-0.01	-0.13	0.645
APIS-P(R) Test 5	0.03	0.69	0.04	0.05	0.04	-0.03	0.517
APIS-P(R) Test 6	0.04	0.41	0.02	0.10	-0.09	-0.17	0.779
APIS-P(R) Test 7	0.23	0.35	-0.18	0.03	0.02	-0.07	0.789
APIS-P(R) Test 8	0.12	0.43	-0.01	-0.04	-0.03	-0.05	0.794
B19 Error time	-0.24	-0.09	0.07	-0.22	-0.2	-0.1	0.830

B19 Number of errors	-0.24	-0.18	0.06	-0.23	-0.13	-0.2	0.798
IVE Impulsivity	0.03	0.12	-0.03	0.00	0.09	-0.37	0.839
IVE Empathy	-0.04	0.24	0.02	0.00	0.00	0.34	0.822
IVE Venturesomeness	-0.05	0.25	-0.03	0.08	0.03	0.0	0.927

The results presented indicate an acceptable model fit (*RMSEA=0*.084, *TFI=0*.786), although the proportion of variance explained jointly by all factors is only moderately high (0.44). The poor fit of the model to the personality variables is also apparent, with all but the agreeableness scale having very high unique values. This may be due in part to the high homogeneity of the sample across these indicators.

Of the remaining indicators, similarly poor match are presented by the number of correct responses in the *COG s1* test (which was an expected result given the small range of this indicator; the response time on this test interacts well with the factors highlighted); the number of correct responses on the first interval of the *DT s5* test (reason similar to the *COG s1* test; the other *DT s5* test indicators present good parameters; the *APIS-P(R)* battery subtest scores (the overall score presents good parameters) and, somewhat surprisingly, the *B19* test indicators.

Based on factor loadings, one may be tempted to interpret the extracted constructs as follows:

Factor 1 – based primarily on *CORSI s1* test indicators, with some association with visual and spatial test scores in the *APIS-P(R)*battery, reaction speed in the *COG s1* test and in the fastest interval of the *DT s5* test, and motor coordination in the *B19* test; may be interpreted as **visual memory and attentional resources** for effective performance in multitasking situations.

Factor 2 – based primarily on the total score and subtests of the *APIS-P(R)* battery, can be referred to as **general intellectual ability**; this is the factor that explains the largest percentage of the overall score variance.

Factor 3 – based mainly on *COG s1* test indicators and to a small extent on *CORSI s1* test indicator, however, unlike factor 1 it does not cover the psychomotor area; it can be interpreted as **visual processing speed** (perceptual ability).

Factor 4 – based inclusive on speed on the *COG s1* test, number of correct and incorrect responses on the *DT s5* test, and *B19* test indicators; it can be interpreted as **psychomotor speed and coordination**.

Factor 5 – based solely on the number of errors in the *DT s5* test, is the least clear-cut for interpretation due to the generally poor performance of these indicators. It can be cautiously interpreted as **impulsivity (poor response control) at the motor level**.

Factor 6 – based on some indicators from the personality structure (mainly neuroticism, agreeableness, conscientiousness and empathy) and the first subtest from the *APIS-P(R)* battery can be interpreted as **an indicator of emotional and social maturity**.

Predictive validity of factors from factor analysis

The final step in the analysis was to see how well the previously highlighted factors allow for predicting whether, three years from the admission tests to OLL, a candidate was likely to pass the psychological examination for admission to LAW. Because the predicted variable had the zero-one form, a logistic regression model was used. The results of the analysis are presented below in (Table 4):

Table 4. A logistic regression model predicting the positive result of admission test to LAW based on scores from the admission test to OLL (presented based on factors).

Predictor	Regression parameter	Standard error	z	p
Constant	1.208	0.271	4.454	< 0.001
Factor 1	0.225	0.252	0.893	0.372
Factor 2	0.589	0.273	2.158	0.031
Factor 3	0.23	0.256	0.897	0.370
Factor 4	0.487	0.250	1.945	0.052

Factor 5	-0.392	0.246	-1.592	0.111
Factor 6	0.391	0.256	1.527	0.127

DISCUSSION

The analysis presented above shows that the tests of cognitive and psychomotor ability used in admission tests to OLL allow us to predict with some confidence the level of functioning of candidates at the end of high school, when testing candidates for LAW. This result is consistent with studies available in the literature that also indicate a significant (relative to, for example, personality traits) contribution of cognitive and psychomotor abilities to the assessment of aviation personnel competency [19]. At the same time, most of the tests show developmental changes in this group in the area of response control (the effect size indicator for the Wilcoxon's r difference above 0.5) – this is particularly visible in the area of indicators based on response time and reflecting resistance to difficult situations and performance in situations of prolonged fatigue. This is an overwhelmingly positive result considering that psychomotor control and coordination is an area undergoing intensive training at a young age, which potentially could have made it difficult to predict outcomes possible several years after the first test. In contrast, visual-memory functions remain relatively stable (Wilcoxon's r equal to 0.36 or less).

Slightly less consistency is evident in the area of personality traits, examined by self-report methods. Comparing the two measures reveals much more variation, with moderate stability appearing only within traits related to sociability and activity. This result is somewhat expected, as the personality profile of OLL candidates – particularly within traits related to emotional maturity – is still forming, making long-term predictions about the test subjects' functioning much more difficult. At the same time, it should also be noted that self-report methods in selection diagnosis have reduced accuracy due to the tendency of some test subjects to overly positive self-presentation [7], which secondarily may have increased the error variance of the results obtained and, as a result, lead to lower stability coefficients.

The least expected result is the low correlation of scores on tests involving visual-spatial abilities. The test performed on LAW candidates correlates moderately only with the test examining abstract-logical abilities, which is reasonable since *PST* scores, in addition to pure spatial abilities, are also dependent on the test subjects' intellectual resources. However, the lack of convergent validity, as expressed by the low correlation with the *APIS-P(R)* subtests based on spatial abilities, may indicate either a mismatch between the construct and the measurement of other aspects of visual-spatial abilities in the testing of OLL and LAW candidates, or large and irregular (i.e., dependent on unsystematic additional factors) developmental changes in the area of these abilities, making it impossible to realistically predict this type of competence at such a young age as middle school graduates.

In general, the entire battery of tests used in testing OLL candidates can be considered a reliable source of information about the expected results obtained during admission tests to LAW. The primary research question, however, was the extent to which this information allows for prediction (several years in advance) of positive qualification to military studies. The logistic regression model constructed in the analyses has limited predictive power, with a pseudo R^2 (Tjur, 2009) of only 0.16. The primary significant predictor is general intellectual ability, which is understandable because not only is it an important (and relatively stable in measurement) component of the assessment, but it also illustrates the ability to learn and acquire skills, which is an important prerequisite for preparation for LAW candidate testing. This result is consistent with reports available in the literature indicating an important role of general intellectual ability in military pilot selection [6,23]. The second most important indicator that was significant at the level of statistical trend was found to be psychomotor speed and coordination. The other predictors did not reach the level of statistical significance. These results show that despite some reproducibility of the results from the first and second tests, the most important factor to predict success in flight training qualification is not the level of special abilities but the potential to learn and acquire skills, confirming the significant role of training in the formation of cognitive abilities in professionally relevant areas [1].

CONCLUSION

The obtained results, although still based on a relatively small sample, allow us to draw preliminary conclusions regarding psychological testing of OLL candidates conducted at WIML. First, the design of the diagnostic battery appears adequate and well aligned with the battery of psychological tests designed for LAW candidates used with older adults. The only exceptions are tests used to measure visual-spatial ability. Second, despite the apparent developmental changes (illustrated, among other things, by significantly better scores in tests, which are repeated in identical form in both batteries), the results obtained when testing candidates for OLL maintain a satisfactory stability (moderate correlations of results after as much as a three-year interval should be considered a very satisfactory result). Third, the greatest predictive power comes from general intellectual ability, which provides a base that enables the acquisition and consolidation of other skills. Contrary to what was predicted, working memory and attention resources play a limited role (and at the same time are characterized by a decent stability), while psychomotor performance has a visible predictive power, and at the same time is characterized by a strong developmental tendency (a large change in results after three years, with moderate stability).

AUTHOR'S DECLARATION

Study Design: Piotr Zieliński. **Data Collection:** Piotr Zieliński. **Manuscript Preparation:** Piotr Zieliński. The Author declares that there is no conflict of interest.

REFERENCES

1. Adler AB, Bliese PD, Pickering MA, Hammermeister J, Williams J, Harada C, Csoka L, Holliday B, & Ohlson C. Mental skills training with basic combat training soldiers: A group-randomized trial. Journal of Applied Psychology 2015; 100(6):1752–1764.
2. ALMamari K, Traynor A. Predictive validity of the Air Force Officer Qualifying Test (AFOQT) for pilot performance: a meta-analytic investigation at the subtest level. Aviation Psychology and Applied Human Factors 2020; 10(2):70-81.
3. Bauer H, Guttmann G, Leodolter M, Leodolter U. Manual for the Sensorimotor Coordination Test. Modling, Austria: G. Schuhfried GmbH; 2002.
4. Bonnardel R. Double labyrinth test. Modling, Austria: G. Schuhfried GmbH; 2003.
5. Costa PTJr, McCrae R. Inwentarz osobowości NEO Wersja Zrewidowana (NEO-PI-R) and Pięcioczynnikowy Inwentarz NEO (NEO-FFI). Podręcznik dla profesjonalistów. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego; 2010.
6. de Kock F, Schlechter A. Fluid intelligence and spatial reasoning as predictors of pilot training performance in the South African Air Force (SAAF). SA Journal of Industrial Psychology 2009; 35(1):31-38.
7. Galić Z, Jerneić Ž, Kovačić MP. Do applicants fake their personality questionnaire responses and how successful are their attempts? A case of military pilot cadet selection. International Journal of Selection and Assessment 2012; 20(2):229-241.
8. Gordon HW, Leighty R. Importance of specialized cognitive function in the selection

of military pilots. Journal of Applied Psychology 1988; 73(1):38-45.

9. Griffin GR, Koonce JM. Review of psychomotor skills in pilot selection research of the U.S. military services. The International Journal of Aviation Psychology 1996; 6(2):125-147.
10. Grössenbrunner P, Neuwirth W. Pilot's Spatial Test. Release 22.00. Mödling: Dr.G.Schuhfried Ges.m.b.H; 2002.
11. Hertzog C, Kramer AF, Wilson RS, Lindenberger U. Enrichment Effects on Adult Cognitive Development: Can the Functional Capacity of Older Adults Be Preserved and Enhanced? Psychol Sci Public Interest. 2008; 9(1):1-65.
12. Horn J. A rationale test and test for the number of factors in factor analysis. Psychometrika 1966; 30:179-185.
13. Hunter DR, Burke EF. Predicting aircraft pilot-training success: a meta-analysis of published research. The International Journal of Aviation psychology 1994; 4(4):297-313.
14. IATA; International Air Transport Association. Guidance Material and Best Practices for Pilot Aptitude Testing (2nd ed.). Montreal-Geneva: IATA; 2012.
15. JAA; Join Aviation Authorities. JAA Manual of Civil Aviation Medicine. Colorado: GED; 2009.
16. Jaworowska A. IVE. Kwestionariusz impulsywności. Impulsywność. Skłonność do ryzyka. Empatia. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego; 2011.
17. Jaworowska A, Szustrowa T. Test matryc Ravena w Wersji dla Zaawansowanych. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego; 1991.
18. King RE, Carreta TR, Retzlaff P, Barto E, Ree MJ, Teachout MS. Standard cognitive psychological tests predict military pilot training outcomes. Aviation Psychology and Applied Human Factors 2013; 3(1):28-37.
19. Martinussen M. Psychological Measures as predictors of pilot performance: a meta-analysis. The International Journal of Aviation Psychology 1996; 6(1):1-20.
20. Matczak A, Jaworowska A, Ciechanowicz A, Stańczak J, Zalewska E. Bateria testów

APIS-P(R). Podręcznik. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego; 2005.

21. Neuwirth W, Benesch M.. DT. Determinationstest. Version 32.00. Mödling: Schuhfried GmbH; 2007.
22. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2017.
23. Ree MM, Carretta TR. Central role of g in military pilot selection. The International Journal of Aviation Psychology 1996; 6(2):111-123.
24. Revelle W. psych: procedures for psychological, psychometric, and personality research. R package version 1.7.5; 2017.
25. Revelle W, Rocklin T. Very Simple Structure: an alternative procedure for estimating the optimal number of interpretable factors. Multivariate Behavioral Research 1979; 14:403-414.
26. Schellig D. Corsi Block-Tapping test. Version 22.00. Mödling: Dr.G.Schuhfried Ges.m.b.H; 2007.
27. Schmiedek F, Lövdén M, & Lindenberger U. Younger adults show long-term effects of cognitive training on broad cognitive abilities over 2 years. Developmental Psychology 2014; 50(9):2304–2310.
28. Siuta J. Inwentarz osobowości NEO-PI-R Paula T. Costy Jr i Roberta R. McCrae. Adaptacja Polska. Podręcznik. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego; 2006.
29. Spinner B. Predicting success in primary flying school from the Canadian automated pilot selection system: derivation and cross-validation. The International Journal of Aviation Psychology 1991; 1(2):163-180.
30. Tjur T. Coefficients of determination in logistic regression models – a new proposal: the coefficient of discrimination. The American Statistician 2009; 63:366-372.
31. Velicer W. Determining the number of components from the matrix of partial correlations. Psychometika 1976; 41:321-327.
32. Wagner M, Karner T. Cognitrone. Version 37.00. Mödling: Schuhfried GmbH; 2006.

33. Zawadzki B, Strelau J, Szczepaniak P, Śliwińska M. NEO-FFI. Inwentarz osobowości Paula T. Costy Jr i Roberta R. McCrae. Adaptacja Polska. Podręcznik. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego; 1995.