

# Ab Initio Personality Testing For Pilots: The Maltese Context

Katya S. De Giovanni<sup>1</sup>

<sup>1</sup>Department of Psychology, Faculty for Social Wellbeing, University of Malta, Malta

**Abstract. Introduction.** The German Wings incident in March 2015 has focused attention on the mental wellbeing of pilots. Legislation by the European Union Aviation Safety Agency (2018) requires a psychological assessment with the Maltese Civil Aviation Directorate fully implementing this regulation. **Research question.** Are the norms on the NEO PI-R for the Maltese Ab Initio Pilot Population significantly different to those for Aviation personnel on the Revised NEO Personality Inventory Manual (Costa & McCrae, 2006)? **Method.** A sample of N=55 Maltese Pilots (Male=49; Female=6) were assessed. Scores were compared using the IBM SPSS Version 26 Software. **Results** indicated significant differences across most scores. **Discussion.** This creates a number of issues for the Maltese Aviation psychologists in practice.

**Keywords:** Malta, Personality Testing, Five-Factor Theory, NEO PI-R, Aviation.

## Introduction

The suicide on the German Wings flight 9525 in March 2015 has focused attention on the mental wellbeing of pilots and how safety can be enhanced through more rigorous selection and support processes by airlines in recruiting pilots. Cooper, White and Lauber's (1979) 10 year review as reported in Chidester et al. (1991) indicates that accident resulting from a lack of knowledge or technical skill were extremely rare and instead breakdowns in communication and workload distribution were typically cited as causal factors. Legislation by the European Union Aviation Safety Agency (2018) has required a psychological assessment to be carried out and the Maltese Civil Aviation Directorate has fully implemented this requirement.

However, the most recent literature (ALMamari & Traynor, 2019) has focused on the important role of Test Battery composite scores in selection and assessment. In another study by ALMamari and Traynor (2020), results from the Air Force Officer Qualifying Test (AFOQT) indicated that pilot performance is best predicted by subtests indicative of perceptual speed, aviation-related aptitude and knowledge, and quantitative ability constructs, and least predicted by subtests indicative of verbal ability construct. Recent research has thus shown that these test batteries do not normally include personality assessments but are based on tests which examine skills in relation to flying. Personality can be described as the 'dynamic and organized set of characteristics possessed by a person that uniquely influences his or her cognitions, motivations and behaviours in various situations.' (Ryckman, 2013 as cited in Gao & Kong, 2016, pg.1). Callister et al. (1997) state that despite the controversy over the relationship between "normal" personality characteristics and pilot performance, there is little argument that there are "abnormal" personality characteristics which are undesirable as highly anxious, hostile, or impulsive pilots probably should not fly. The literature on personality testing with regards to pilot assessment both civil and military is scant and mixed and primarily carried out in the military environment (Gao & Kong, 2016). By citing (Hunter & Burke, 1995, p. 46) King (2014, p. 61) attempts to clarify and cultivate the "thorny issue of personality assessment". Johnston (1996) states that although the psychological testing of applicants for pilot training and employment is now a well-established and widely accepted activity in most parts of the world, it is only recently that psychological testing has been potentially envisaged to have a substantive role in pilot testing. Furthermore, Johnston (1996) states that psychological tests in pilot selection seem to be well established, but that the relative merits and validities of particular

tests, or batteries of tests, frequently are debated and challenged. Gao and Kong (2016) purport that besides the need to better understand psychological traits of the pilot population, studies of pilots' personality types usually carry the mission of assisting pilot selection and predicting future training success. Amongst these studies researchers have deemed the NEO-PI-R to be suitable in the facilitation of pilot selection. (Fitzgibbons et al., 2004). Chidester et al. (1991) indicate that although a great deal of effort within the aviation community has been focused on ensuring technical expertise and new efforts highlight attitudes associated with crew coordination, personality factors have been relatively unexplored. Past failures to find linkages between personality and performance were due to a combination of inadequate statistical modeling, premature performance evaluation and reliance on data gathered in contrived as opposed to realistic situations. An extensive meta-analysis of the literature over the past 20 years regarding military pilot selection conducted by Paullin et al. (2006) reported that personality traits relevant to pilot performance include conscientiousness, integrity, achievement orientation, emotional stability, resilience, cognitive flexibility, openness, self-confidence, self-esteem, and risk tolerance (Chappelle et al., 2010). There are personality traits that represent the "wrong stuff" such as high anxiety, hostility, depression, isolation, and impulsivity. These traits can conceivably elevate the risk for an aviation mishap where the threat to human life is high. On the other hand, the "Right Stuff" is mainly characterized by high levels of Conscientiousness and Agreeableness as well as low Neuroticism for the NEO PI-R (Musson et al. as cited in Mittelstadt et al., 2016)

The NEO PI-R assesses the Big Five Factors of Personality, namely Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness. Mittelstadt et al. (2016) mention that the NEO PI-R is popular because of its claim to assess a comprehensive personality profile and its relative simplicity of producing scores on five broad factors. Judge et al. (2013) however mention that the lower order traits like the facets incorporated into the five domains are also important. When using the domains alone, lower order criterion-related validities might be masked out because of the differential relationships of the facets that are included in that factor. In Aviation, low levels of Neuroticism and high levels of Extraversion were found to be characteristic for commercial airline pilots and predictive for military aviation training outcome. Martinussen and Torjussen (1998) state that some tests are better predictors of pilot performance than others. The best predictors of pilot performance are tests measuring certain cognitive and psychomotor abilities and biographical information whereas tests measuring personality and general intelligence have lower predictive validity. (Hunter & Burke and Martinussen, as cited in Martinussen & Torjussen, 1998).

Gao and Kong (2016) point out that from their review of literature, it can be seen that much of the past research surrounding personality and pilots mostly focus on military and civil commercial operations with very few looking at the civil ab-initio flight training environment. Furthermore, to date, there have been no such studies pertaining to the Maltese Context. More so there have been no studies which have compared Maltese Mean Scores to the Revised Neo Personality Norm Scores. Nevertheless, several studies demonstrate that pilots as a group differ from the general population and for this reason experienced aviation psychologists use pilot based normative data whenever this is possible. (Retslaff & Gilbertini, as cited in Callister et al. 1997). This paper is an attempt at providing such insights.

## **Method**

The scores from the administration of the NEO PI-R to Maltese individuals aiming to follow a pilot course were recorded. The administration of the psychometric test was carried out by the same qualified organizational and aviation psychologist on a one-to-one basis in a private Aeromedical Health Screening Clinic. The test was administered in standard English and any

queries with regards to the terms were always given the same meaning in the Maltese language. Results from the scores were anonymized and inputted into the IBM SPSS Version 26. The mean and standard deviation of both the domain and the facet scores were reported. A paired sample T-Test comparing the means of the Maltese Sample with the Aviation (N=45) sample reported in the manual of the Revised NEO Personality Inventory (2006).

### Publication ethics

Ethical Clearance was filed with the University of Malta Faculty for Social Wellbeing Research Ethics Committee. Since the data were analysed as a group and there was no means of identifying any scores on an individual basis, an application was filed for auditing purposes. There were no queries and the study could proceed.

## Results

The results will be reported per domain and facet scores. The results clearly show significant differences between the Maltese sample and the NEO sample for Aviation Personnel on the manual of the Revised Neo Personality Inventory (2006). Although the full results are reported in the tables, only results with significant differences at  $P < 0.05$  will be reported in text.

Table 1: Paired Samples Test for Neuroticism Domain and Facet Scores

		Paired Differences			df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean		
Pair 1	AnxietyM - AnxietyNEO	-.87273	8.52416	1.14940	54	.451
Pair 2	AngryHM - AngryHNEO	2.07273	8.31967	1.12182	54	.070
Pair 3	DepressionM - DepressionNEO	2.84545	6.26045	.84416	54	.001
Pair 4	SelfConsciousnessM - SelfConsciousnessNEO	4.26364	7.69984	1.03825	54	.000
Pair 5	ImpulsivenessM - ImpulsivenessNEO	-2.82727	6.94703	.93674	54	.004
Pair 6	VulnerabilityM - VulnerabilityNEO	1.88182	8.96514	1.20886	54	.125
Pair 7	NeuroticismM - NeuroticismNEO	1.69091	7.60573	1.02556	54	.105

As shown in Table 1 there are significant differences on three of the facet scores. For Depression ( $P=0.001$ ), the Maltese Sample mean (52.75) was significantly higher than the NEO Mean (49.90). For Self-Consciousness ( $P=0.000$ ), the Maltese sample mean (55.16) was significantly higher than the NEO Mean (50.9). In the case of Impulsiveness ( $P=0.004$ ), the Maltese sample (47.67) was significantly lower than the NEO Sample (50.50). There was no significant difference between samples on the overall Domain Score of Neuroticism.

Table 2: Paired Samples Test for Extraversion Domain and Facet Scores

		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Pair 1	WarmthM - WarmthNEO	.34545	9.64407	.266	54	.792
Pair 2	GregariousnessM - GregariousnessNEO	1.20000	10.52158	.846	54	.401
Pair 3	AssertivenessM - AssertivenessNEO	-4.49091	9.82037	-3.391	54	.001
Pair 4	ActivityM - ActivityNEO	-12.80000	8.02542	-11.828	54	.000
Pair 5	ExcitementSeekingM - ExcitementSeekingNEO	1.76364	7.80732	1.675	54	.100
Pair 6	PositiveEmotionM - PositiveEmotionNEO	-2.41818	7.16134	-2.504	54	.015
Pair 7	ExtraversionM - ExtraversionNEO	-2.90909	9.38137	-2.300	54	.025

As shown in Table 2 there are significant differences on three of the facet scores. For Assertiveness ( $P=0.001$ ), the Maltese Sample mean (48.30) was significantly lower than the NEO Mean (52.80). For Activity ( $P=0.000$ ), the Maltese sample mean (43.00) was significantly lower than the NEO Mean (55.80). In the case of Positive Emotion ( $P=0.000$ ), the Maltese sample (50.58) was significantly lower than the NEO Sample (53). There was also a significant difference on the overall Domain Score of Extraversion where the Maltese sample (49.09) was significantly lower than the NEO Sample (52.00).

Table 3: Paired Samples Test for Openness Domain and Facet Scores

		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Pair 1	FantasyM - FantasyNEO	-.73636	7.51779	-.726	54	.471
Pair 2	AestheticsM - AestheticsNEO	-.48182	7.83178	-.456	54	.650
Pair 3	FeelingsM - FeelingsNEO	-7.93636	7.39360	-7.961	54	.000
Pair 4	ActionsM - ActionsNEO	-5.10000	6.61816	-5.715	54	.000
Pair 5	IdeasM - IdeasNEO	3.47273	9.64721	2.670	54	.010
Pair 6	ValuesM - ValuesNEO	-10.40909	8.14196	-9.481	54	.000
Pair 7	OpennessM - OpennessNEO	-4.36364	6.22945	-5.195	54	.000

As can be seen in Table 3 there are significant differences on four of the facet scores. For Feelings ( $P=0.000$ ), the Maltese Sample mean (43.96) was significantly lower than the NEO Mean (51.90). For Actions ( $P=0.000$ ), the Maltese sample mean (44.60) was significantly lower than the NEO Mean (49.70). In the case of Ideas ( $P=0.010$ ), the Maltese sample (51.47) was significantly higher than the NEO Sample (48). Furthermore, there was a significant difference ( $P=0.000$ ) between the Maltese Sample (40.49) and the NEO Sample (50.90) on the facet score of Values. There was also a significant difference on the overall Domain Score of Openness ( $P=0.000$ ) where the Maltese sample (43.83) was significantly lower than the NEO Sample (48.20).

Table 4: Paired Samples Test for Agreeableness Domain and Facet Scores

		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Pair 1	TrustM - TrustNEO	-13.97273	8.15102	-12.713	54	.000
Pair 2	StraightforwardnessM – Straightforwardness NEO	4.01818	8.60984	3.461	54	.001
Pair 3	AltruismM - AltruismNEO	3.24545	11.01921	2.184	54	.033
Pair 4	ComplianceM - ComplianceNEO	.37273	9.39475	.294	54	.770
Pair 5	ModestyM - ModestyNEO	-.14545	9.49623	-.114	54	.910
Pair 6	TenderMindednessM – Tendermindedness NEO	7.34545	9.75101	5.587	54	.000
Pair 7	AgreeablenessM - AgreeablenessNEO	.00909	8.92120	.008	54	.994

As can be seen in Table 4 there are significant differences on four of the facet scores. For Trust ( $P=0.000$ ), the Maltese Sample mean (41.93) was significantly lower than the NEO Mean (55.90). For Straightforwardness ( $P=0.001$ ), the Maltese sample mean (57.62) was significantly higher than the NEO Mean (53.60). In the case of Altruism ( $P=0.033$ ), the Maltese sample (53.95) was significantly higher than the NEO Sample (50.70). Furthermore, there was a significant difference ( $P=0.000$ ) between the Maltese Sample (57.34) and the NEO Sample (50.00) on the facet score of Tendermindedness. There was no significant difference on the overall Agreeableness Domain Score.

Table 5: Paired Samples Test for Conscientiousness Domain and Facet Scores

		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Pair 1	CompetenceM - CompetenceNEO	-.78182	9.33827	-.621	54	.537
Pair 2	OrderM - OrderNEO	5.89091	7.59062	5.756	54	.000
Pair 3	DutifulnessM - DutifulnessNEO	1.83636	8.77447	1.552	54	.126
Pair 4	AchievementStrivingM - AchievementStrivingNEO	-8.34545	8.75218	-7.072	54	.000
Pair 5	SelfDisciplineM - SelfDisciplineNEO	3.16364	9.99252	2.348	54	.023
Pair 6	DeliberationM - DeliberationNEO	12.18182	10.88916	8.297	54	.000
Pair 7	ConscientiousnessM - ConscientiousnessNEO	5.51818	10.25001	3.993	54	.000

As can be seen in Table 5 there are significant differences on four of the facet scores. For Order (P=0.000), the Maltese Sample mean (55.29) was significantly higher than the NEO Mean (49.40). For Achievement Striving (P=0.000), the Maltese sample mean (50.65) was significantly lower than the NEO Mean (59.00). In the case of Self-Discipline (P=0.023), the Maltese sample (52.96) was significantly higher than the NEO Sample (49.80). Furthermore, there was a significant difference (P=0.000) between the Maltese Sample (61.38) and the NEO Sample (49.2) on the facet score of Deliberation. There was also a significant difference on the overall Conscientiousness Domain Score (P=0.000) where the Maltese sample has a mean of 54.41 and the NEO sample had a mean of 48.90.

## Discussion

Literature has focused mainly on the differences between males and females and not really on ethnic differences as was the scope of this study. Furthermore, most compared findings to the general population and not an aviation norm group as was the scope of this study.

Chappelle (2010) indicated that on scores on the NEO PI-R USAF female pilots are more extroverted on than females in the general population. They are more interpersonally gregarious, assertive, outgoing, excitement seeking and expressive of positive emotions. As a group they are more open to new experiences, inner feelings and emotions and are more willing to consider new and perhaps unconventional ideas. Female pilots were inclined to be tough minded, straightforward, proud, competitive, achievement oriented and self-confident. Furthermore, they were more open to inner feelings and emotions while also being more capable of handling stress and remaining composed in difficult and highly challenging situations. Callister (1999) in his study using the NEO personality Inventory over 1301 US Airforce Pilots concluded that the average male pilot was altruistic yet highly competitive, sceptical and tough minded. Compared with the general male adult norms it was noted that the male student pilot sample had greater levels of extraversion and lower levels of agreeableness. In this study it was note that female students showed very similar characteristics to the male

sample that is they showed higher levels of extraversion and openness and lower levels of agreeableness in comparison to the general population.

Fitzgibbons et al., (2004) state that the basic pilot personality profile is of an emotionally stable individual who is low in anxiety, vulnerability, angry hostility, impulsiveness and depression. Pilots also tend to be very conscientious and are high in deliberation, achievement striving, competence and dutifulness. Pilots also tend to be trusting and straightforward as well as active with a high level of assertiveness. Wakcher et al. (2003) point out that the person that is drawn to the occupation of airline pilot is substantially more reserved, intelligent, emotionally stable, dominant, enthusiastic, conscientious, bold, trusting, self-assured, conservative, socially precise and relaxed than is the general population. The Maltese Sample of ab initio pilots investigated for this paper is significantly more depressed, self-conscious, less impulsive, less assertive, less active and displayed less positive emotion than the norms on the aviation sample on the Manual of the Revised Neo Personality Inventory. The Maltese Sample overall displays less extraversion as a Domain score and also attained a significantly lower score in feelings and actions. Wood et al. (2015) state that interpersonal skills play a role in every occupation and are known to be an important characteristic of successful pilots together with their above average level of extraversion. Moreover, a lower score was attained on values and Openness in general as a Domain Score. Furthermore, the Maltese Sample attained significantly lower scores on trust and achievement striving facet scores. On the other hand, the Maltese scored significantly higher on ideas, straightforwardness, altruism, tendermindedness as well as Conscientiousness as a Domain score. Conscientiousness means that pilot students have a strong desire to do well and are highly efficient and organized (Gao & Kong, 2016). Fitzgibbons et al., (2004) mention that pilots tend to have personality domains and dimensions that are consistent across background and circumstance. This paper now questions this assertion given the results attained from this study.

The current study has its limitations. First of all, all assessments were carried out by the same psychologist and due to limitations of resources, there was no supervision being carried out on the employed practice. Moreover, at times the English used in the was cumbersome to understand. Although the same translation and examples were always given, these could have impinged on the standardization of the instrument. The sample size was also not very big although comparable to the norm sample for Aviation personnel used in the Revised NEO Personality Inventory. Moreover, as Martinussen and Torjussen (2009) point out, for smaller nations it is difficult to conduct validation studies with large enough samples to ensure adequate statistical power. One also needs to point out that the norm sample of Aviation personnel was made up by both pilots and also air-traffic controllers and this could explain the significant differences which were elicited from this study. One should however notice that the five-factor model represents the results of a growing consensus that five dimensions adequately describe the personality domain. (Siem, 1992).

Given the above results, one queries the relevance of using the Norm scores as stipulated by the Revised NEO Personality Inventory Manual (2006) and the current practice which is persistent on the Maltese Islands in this regard. It is hoped that this paper will shed light on possible shortcomings of the use of the Revised NEO Personality Inventory and also other Personality Inventories as used in other countries. Undoubtedly there needs to be further studies in this regard both on the Maltese Islands possibly also investigating gender differences. A valid psychological assessment is a crucial part of the process of providing psychological services in unique and high-risk positions. Aside from ability and motivation for military flying, personality characteristics play a key role in succeeding as a pilot (Chappelle, 2010).

Callister et al. (1997) state that pilots are valuable personnel assets who must be provided the highest level of psychological services. Therefore, valid psychological assessment is a critical step in this process. Large studies using reliable, valid and clinically relevant tests can yield the type of data necessary to improve the psychological service available to the flying community (Callister et al., 1997). Further research could follow up on the sample tested for this study and have a retake of the same instrument namely the Revised NEO Personality Inventory (2006) to see whether there were any changes due to organizational socialization and training. Literature seems to support the fact that individual traits do tend to vary throughout the adult life due to maturation and social factors although the rank ordering remains fairly stable over spans of up to 45 years (Conley as cited in Fitzgibbons et al., 2004)

## References

- ALMammari, K. & Traynor, A. (2019). Multiple test batteries as predictors for pilot performance: A meta-analytic investigation. *International Journal of Selection and Assessment*. 27, 337-356.
- ALMammari, K. & Traynor, A. (2020). Predictive Validity of the Air Force Officer Qualifying Test (AFOQT) for Pilot Performance. *Applied Psychology and Human Factors*. 10(2), 70–81.
- Callister, J. D., King, R. E., Retzlaff, P. D. & Marsh, R. W. (1997) Using the NEO-PI-R to assess the personality of US Air Force Pilots. *United States Air Force Armstrong Laboratory*. Retrieved from: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a328908.pdf>
- Callister, J. (1999). Revised NEO personality inventory profiles of male and female US Air Force Pilots. *Military Medicine*, 164(12), 885-890. Retrieved from <https://search-proquest-com.ejournals.um.edu.mt/docview/217042744?accountid=27934>
- Chappelle, W. L., Novy, P. L., Sowin, T. W. & Thompson, W. T. (2010). NEO PI\_R Normative Personality Data That Distinguish U.S. Air Force Female Pilots. *Military Psychology*. 22, 158-175.
- Chidester, T.R., Helmreich, R. L., Gregorich, S. E., & Geis, C. E. (1991). Pilot Personality and Crew Coordination: Implications for Training and Selection. *The International Journal of Aviation Psychology* 1(1), 25-44.
- Costa, P. T. & McCrae, R. R. (2006). *Revised Neo Personality Inventory (NEO PI-R) Manual*. Hogrefe: Oxford.
- Fitzgibbons, A., Schutte, P. & Davis, D. (2004). Pilot personality profile using NEO-PI-R. *National Aeronautics and Space Administration*. Retrieved from <https://ntrs.nasa.gov/citations/20040191539>
- Gao, Y. & Kong, S. (2016). Personality types of pilot students: A study of an Australian collegiate aviation program. *International Journal of Aviation, Aeronautics and Aerospace*. 3(3)6.
- Judge, T. A., Rodell, J. B., Klinger, R. L. Simon, L. S. Crawford, E. R. (2013). Hierarchical Representations of the five factor model of personality in predicting job performance. *Journal of Applied Psychology*. 98(6), 875-925.
- King, R. E. (2014) Personality (and Psychopathology Assessment in the Selection of Pilots. *The International Journal of Aviation Psychology* 24(1) 61-73.
- Martinussen, M. & Torjussen, T. (2009) Pilot Selection in the Norwegian Air Force: A Validation and Meta-Analysis of the Test Battery. *The International Journal of Aviation Psychology*. 8(1), 33-45.
- Mittelstadt, J. M., Pecena, Y., Oubaid, V. & Maschke, P. (2016). Psychometric Personality Differences Between Candidates in Astronaut Selection. *Aerospace Medicine and Human Performance* 87(11), 933-939.
- Siem, F. M. (1992). Predictive Validity of an Automated Personality Inventory for Air Force Pilot Selection. *The International Journal of Aviation Psychology*. 2(4), 261-270.

- Wakcher, S., Cross, K. & Blackman, M. C. (2003). Personality Comparison of Airline Pilot Invuments, Applicants and the Geenal Population Norms on the 16PF. *Psychological Reports* 92, 773-780.
- Wood, J., Shurlow, C. & Haynes, J. (2015). Objective versus Subjective Military Pilot Selection Methods in the United States of America. *Air Force Research Laboratory*. Retrieved from: <https://apps.dtic.mil/sti/pdfs/ADA626003.pdf>

## Contact Information

Dr Katya De Giovanni  
Department of Psychology  
Faculty for Social Wellbeing  
University of Malta  
Malta.

katya.degiovanni@um.edu.mt