Psychometric Personality Differences Between Candidates in Astronaut Selection

Justin M. Mittelstädt; Yvonne Pecena; Viktor Oubaid; Peter Maschke

INTRODUCTION: This paper investigates personality traits as potential factors for success in an astronaut selection by comparing personality profiles of unsuccessful and successful astronaut candidates in different phases of the ESA selection procedure. It is further addressed whether personality traits could predict an overall assessment rating at the end of the selection.

METHODS: In 2008/2009, ESA performed an astronaut selection with 902 candidates who were either psychologically recommended for mission training ($N = 46$) or failed in basic aptitude ($N = 710$) or Assessment Center and interview testing ($N = 146$). Candidates completed the Temperament Structure Scales (TSS) and the NEO Personality Inventory Revised (NEO-PI-R).

RESULTS: Those candidates who failed in basic aptitude testing showed higher levels of Neuroticism ($M = 49.8$) than the candidates who passed that phase ($M = 45.4$ and $M = 41.6$). Additionally, candidates who failed in basic testing had lower levels of Agreeableness ($M = 132.9$) than recommended candidates ($M = 138.1$). TSS scales for Achievement ($r = 0.19$) and Vitality ($r = 0.18$) showed a significant correlation with the overall assessment rating given by a panel board after a final interview.

DISCUSSION: Results indicate that a personality profile similar to Helmreich's "Right Stuff" is beneficial in astronaut selection. Influences of test anxiety on performance are discussed.

KEYWORDS: astronaut selection, personality assessment, Temperament Structure Scales, Big Five.

Mittelstädt JM, Pecena Y, Oubaid V, Maschke P. *Psychometric personality differences between candidates in astronaut selection*. Aerosp Med Hum Perform. 2016; 87(11):933–939.

The importance of personality assessment in the selection of astronauts becomes apparent when considering the impact of personality on communication, group cohesion, and overall interpersonal effectiveness.[15] During long-duration missions to the ISS, personality becomes ever more important, as the relevance of personality increases with the duration.[29] As a result, researchers recommend including psychological tests to assess personality and interpersonal skills for the selection of astronauts.[13,14,27]

For psychometric personality assessment in the context of aerospace personnel, different nonpathological personality questionnaires are used to assess the general adult personality. Frequently used in this context is the Personality Characteristic Inventory (PCI) developed by Helmreich.[10] The PCI consists of two broad dimensions: Instrumentality, which is a goal-oriented achievement motivation, and Expressivity, which describes interpersonal capacities and orientations, each with positive and negative aspects. Along with Instrumentality and Expressivity, five more scales were combined in the PCI: Work Orientation, Competitiveness, Mastery, Achievement Striving, and Impatience/Irritability.

Another personality inventory that is often used in the personality assessment of aerospace personnel is the NEO Personality Inventory Revised (NEO-PI-R) or the shortened version, the NEO Five-Factor Inventory (NEO-FFI).[4] Both instruments assess the Big Five factors of personality: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness.

From the German Aerospace Center (DLR), Hamburg, Germany.

This manuscript was received for review in December 2015. It was accepted for publication in July 2016.

Address correspondence to: Justin Mittelstädt, Sportallee 54a, 22335 Hamburg, Germany; Justin.Mittelstaedt@dlr.de.

DOI: 10.3357/AMHP.4548.2016

The NEO-PI-R is popular as a result of its claim to assess a comprehensive personality profile and its relative simplicity of producing scores on five broad factors. Recent findings,[12] however, emphasize the importance of lower order traits like the facets incorporated into the NEO-PI-R (six per factor) or other approaches like the Big Five aspects.[5] When using the Big Five factors alone, lower order criterion-related validities might be masked or cancelled out because of differential relationships of facets that are included in that factor. This is apparent, for example, in Conscientiousness, a Big Five factor that is divided into facets of Industriousness, which is described as achievement-orientation, self-discipline, and purposefulness, and Orderliness, described as deliberation, tidiness, and caution, in the Big Five aspects framework by DeYoung et al.[5] It is easily imaginable that those two aspects have differential validities for criteria of job success. Therefore, it is advisable to expand the analysis of the NEO-PI-R to the facet level or to distinguish between aspects of a factor that have differential validities.

A third instrument for personality assessment are the Temperament Structure Scales (TSS), which were designed and constructed in an aerospace selection environment at DLR[16] and have successfully been used in the selection of pilots,[11] air traffic controllers,[25] and were also used in all three ESA astronaut selection campaigns since 1977.[17,29]

The PCI scales are often associated with the term "Right Stuff," which is used as a synonym for an ideal profile in popular as well as in research literature on astronauts and aviation and is defined by NASA with characteristics beyond what one can assess with a personality questionnaire (e.g., technical aptitude, ability to tolerate isolation, cultural sensitivity).[29] In terms of the PCI scales, the "Right Stuff" is mainly characterized by high levels of Instrumentality and Expressivity and low levels of Negative Instrumentality and Verbal Aggressiveness.[3,22] For the NEO-PI-R, mostly high levels of Conscientiousness and Agreeableness as well as low Neuroticism are associated with the "Right Stuff" profile.[22]

The "Right Stuff" personality profile does indeed show some validity in the aerospace context or environments that are associated with aerospace research. High Instrumentality and Expressivity were related to emotional stability during Antarctic missions.[28] In samples of pilots[3] and astronauts,[18] the "Right Stuff" profile was related to higher ratings on social compatibility and overall job performance.

For the Big Five, relationships with performance in different occupational contexts have been studied. Low levels of Neuroticism and high levels of Extraversion were found to be characteristic for commercial airline pilots[8] and predictive for military aviation training outcome.[1]

In the context of extreme working environments, low Neuroticism was found to be related to emotional stability during an Antarctic mission.[23] Furthermore, high Agreeableness and low Openness enhance professional effectiveness of astronauts as assessed by peer and supervisor ratings.[26] Agreeableness was also found to be beneficial for interpersonal, technical, and leadership peer ratings as well as for supervisor ratings. Openness to Experience was related to lower technical and leadership peer ratings and a lower supervisor rating because as the authors put it: "Openness seems to tap a sort of freefloating intellectual curiosity (. . .)" (p. 914).[26]

For the overall performance, high Agreeableness, low Openness, low Neuroticism for astronauts,[26] and, additionally for Antarctic personnel, low Extraversion and low Conscientiousness[23] have been reported as beneficial. Palinkas et al.[24] identified high Agreeableness, low Neuroticism, low Extraversion and Openness as the second most important indicator for overall performance in isolated and confined environments after motivation and need for achievement.

These studies show different and sometimes contradicting relationships between the Big Five and various parameters of performance, probably due to divergent demands in those occupations. Findings from Antarctic missions are often used for space missions due to the lack of research with astronauts and the relative similarity of both working environments in terms of confinement and hostility. Subjects that engage in either one of these work environments are indeed similar on many different personality scales and especially in their deviation from the normative personality profile in the general population. But astronauts were, in contrast, found to be significantly more achievement-oriented than their counterparts working in Antarctic missions.[21] Therefore, generalizing findings of personality correlates from Antarctic personnel to astronauts are not necessarily false, but have to be done cautiously. The only consistent commonality that most of the studies shared was the negative effect of Neuroticism on different parameters of performance.

The TSS were never linked to the "Right Stuff" profile, but a recent analysis of the construct validity of the TSS in relation to the NEO-PI-R revealed consistent empirical overlaps of both instruments in different samples.[20] This analysis has also shown that the TSS add to the description of personality above the Big Five factors either by giving a more distinctive profile, for example, by distinguishing between Achievement motivation and Rigidity, two components that are combined in the Big Five factor Conscientiousness, or by featuring scales that emphasize aspects that are specifically important for aerospace personnel like the scales Mobility or Vitality.

The TSS have been shown to be predictive for job performance of commercial airline pilots. Those pilots that performed below standard (e.g., had more than one recheck) were emotionally more unstable and aggressive, but also more empathetic.[11] In addition, three scales showed predictive validity toward training criteria in an air traffic controller sample in a recent large scale validation study.[7] Higher Rigidity, lower Aggressiveness, and higher Achievement Motivation corresponded to better training success. So far, there are no published data on relationships of the TSS with performance criteria in an astronaut sample or within the context of an Antarctic mission.

All three personality questionnaires introduced above have been used in different aerospace-related contexts like aviation, extreme environments (Antarctic), or with astronauts and

some scales showed validities for performance in these kinds of environments. But it is questionable whether psychometrically assessed personality traits also influence the selection of astronauts or personnel in extreme environments and a suitability rating prior to the mission.

In previously reported studies on the issue, no differences on the scales of the NEO-PI-R or in PCI clusters were found between successful and unsuccessful astronaut applicants in the final stage of a NASA astronaut selection.[22] Similarly, in the selection of Antarctic personnel, no relationship between the suitability rating after selection interviews and an independently administered personality battery, including the NEO-FFI, were found.[9]

This could in part be explained by the fact that those applicants are typically an already highly selected sample. High psychological and physical demands of space or Antarctic missions can be anticipated; therefore, people who are applying are probably not highly anxious and neurotic. Formal requirements often preselect candidates with high academic or professional success, which often requires a minimum level of conscientiousness as well.

The candidates included in the analysis by Musson et al.[22] were indeed significantly higher on Extraversion, Conscientiousness, and Agreeableness and lower on Neuroticism than a normative comparison sample. These results are supported by Maschke et al.,[17] who found higher mean scores on Openness to Experience, Agreeableness, and Conscientiousness and lower mean scores on Neuroticism for European astronaut candidates even compared to a sample of U.S. Air Force student pilots, another highly selected group.

The special characteristics of astronaut applicant samples may not be the only reason for a lack of relationship with selection success. Although the results by Musson et al.[22] indicate that personality, assessable with a general personality questionnaire, does not differ between successful and unsuccessful astronaut candidates, their results are based on the NASA selection procedure between 1990 and 1994, which strongly focused on a select-out approach of rejecting those candidates with disqualifying psychopathological characteristics and might, therefore, not be applicable to the select-in focused approach used by ESA. The influence of personality traits assessed by a general personality questionnaire might be stronger in a selection procedure that emphasizes those characteristics that suit one for working in space instead of focusing on those characteristics that disqualify one for the job. Additionally, the decisions described in this study are based solely on psychological testing and do not reflect any medical criteria.

This paper investigates personality differences between successful and unsuccessful candidates and between successful candidates and candidates who failed in different phases of the selection procedure in the ESA astronaut selection 2008/2009. Until this point, no systematic analysis of personality differences among astronaut candidates has been conducted using an ESA astronaut sample and with a selection procedure that focuses on a select-in approach.

METHODS

Subjects

Data from $N = 902$ individuals who participated in the first phase of the ESA astronaut selection 2008/2009 were included in this study. The sample consisted of 162 female and 740 male subjects. Their mean age was 33.18 yr with a standard deviation of 3.63 yr. The range was between 24 and 46 yr of age. They originated from 18 different European countries representing all ESA member states at that point in time. All of them completed basic aptitude tests and personality questionnaires. All subjects provided written informed consent before the start of phase 1 testing in which they approved their data being stored and used for evaluation of the applied instruments and for scientific purposes.

After phase 1, a selection based on the results of the basic aptitude testing was administered, leaving 192 candidates, of which 22 were women and 170 men. Compared to phase 1, subjects of phase 2 had a similar age distribution with a mean age of 32.80 and a standard deviation of 3.57 yr. The range decreased to a range between 26 and 45 yr of age.

A total of 46 candidates were finally recommended for further assessment after the psychological examination. These comprised 6 female and 40 male subjects with a mean of 33.48 and a standard deviation of 4.12 yr of age. The age range remained between 26 and 45 yr of age. After a subsequent medical assessment, the remaining candidates were conclusively selected by an ESA internal board.

Materials

Two personality questionnaires were used in order to assess different personality constructs. The TSS[16] are used at DLR in the selection of pilots and air traffic controllers and are specifically designed to assess scales relevant in the context of aerospace applications. It comprises 11 scales [Achievement, Instability, Rigidity, Extraversion, Aggressiveness, Vitality, Dominance, Empathy, Spoiltness (need for luxury), Mobility, Openness], including a scale for social desirability (Openness), with a total of 234 items. TSS answers are given in a dichotomous format either by agreeing or disagreeing to a given statement (yes/no) or by choosing one of two options.

Additionally, the NEO-PI-R[4] was administered. It assesses the Big Five factors of personality using 240 items on a 5-point Likert scale. Each factor consists of six facets with eight items, respectively.

Procedure

In 2008 and 2009, ESA conducted its third astronaut selection to select astronauts for long-duration missions aboard the ISS. The psychological assessment and selection was performed by ESA, and administered by DLR in cooperation with the French Institute for Space Medicine and Physiology (MEDES).

After an online preselection application process, basic aptitude testing was administered in phase 1 of the selection procedure. The choice of basic aptitude tests was based on experience in earlier astronaut selections as well as on experience gained in

the selection of pilots and air traffic controllers. Included in the battery were tests for memory, attention, perception, spatial ability, reasoning, psychomotor coordination, multitasking ability, mental arithmetic, and English skills.[17] Candidates were required to attain at least a certain minimum level in every test. Phase 1 also included the TSS and the NEO-PI-R as personality questionnaires which were not a direct basis for any selection decisions.

Those who were recommended for further testing based on their performance in the aptitude tests were invited to phase 2 testing, with Assessment Center exercises, an individual interview, and a final panel interview.[17] The selection board, consisting of an ESA astronaut, one person from ESA HR personnel, and one psychologist from DLR and MEDES each, respectively, had to conclude with a final overall assessment rating (OAR) between 1 and 9, taking all collected data into account. An OAR of 7 or above indicated a recommendation.

Statistical Analysis

Groups were formed based on the phase that a candidate failed or whether a candidate received a recommendation. Therefore, Group 1 consisted of those candidates that failed in phase 1 (basic aptitude and knowledge tests), Group 2 included candidates that failed in phase 2 (personality and team competencies), and Group 3 contained those candidates that received a recommendation for medical assessment. The assignment of groups by selection phases is illustrated in **Table I**.

Multivariate analyses of variance (MANOVA) were performed between the three groups separately for the Big Five factors, the 11 scales of the TSS, and the 30 Big Five facets. On a significant MANOVA (α-level: 5%), analyses of variance (ANOVA) were performed between the three groups for all scales contained in the significant MANOVA. Based on a significant *P*-value (α-level: 5%, Bonferroni corrected) in the ANOVA, post hoc least significant difference tests with Bonferroni α correction to control for inflated type I error were performed.

In addition, it was taken into consideration whether differences existed between candidates who received a recommendation and those who did not receive a recommendation, which is in line with previous research[22] that distinguished between successful and unsuccessful astronaut candidates and allows for a better comparison with those results. Thus, Groups 1 and 2 combined were tested against Group 3. Furthermore, correlation analyses were conducted using personality data and the OAR ranging from 1 to 9, with 9 indicating the highest degree of suitability. Only candidates recommended for phase 2 testing received a final OAR, so only Group 2 and Group 3 candidates are included in this analysis.

Table I. Analysis Groups by Selection Phases.

	GROUPS FOR STATISTICAL ANALYSES		
SELECTION PHASES	**GROUP 1; *N* = 710**	**GROUP 2; *N* = 146**	**GROUP 3; *N* = 46**
Phase I, *N* = 902	Fail	Pass	Pass
Phase II, *N* = 192	---	Fail	Pass
Recommended, *N* = 46	---	---	Recommended

RESULTS

Means and standard deviations of all personality scales assessed and all groups are shown in **Table II**. As becomes evident, only a few scales differed in their means across the three groups. Two out of three MANOVAs yielded significant results. The MANOVA for the Big Five factors [$F(5,895) = 5.83$; $P < 0.001$; Wilks $\Delta = 0.968$; $\eta^2 = 0.032$] and the Big Five facets [$F(30,871) = 2.76$; $P < 0.001$; Wilks $\Delta = 0.913$; $\eta^2 = 0.089$] were significant, while the MANOVA for the scales of the TSS did not reach significance [$F(10,891) = 1.54$; $P = 0.213$; Wilks $\Delta = 0.985$; $\eta^2 = 0.015$].

For the Big Five factors and facets, additional ANOVAs were performed for all scales with a Bonferroni correction of the α-level. The results of the different one-way ANOVAs are present in **Table III**. Given the overall statistically significant ANOVA, least significant difference tests with Bonferroni correction were performed for all possible group comparisons.

For the Big Five factors of the NEO-PI-R, Neuroticism yielded a significant result in the ANOVA after the Bonferroni correction (Table III). Post hoc tests revealed significant differences between Group 1 and Group 2 ($P_{adjusted} = 0.005$) and between Group 1 and Group 3 ($P_{adjusted} = 0.002$). Although the mean Neuroticism score in Group 3 is lower than in Group 2 (**Fig. 1**), this difference did not reach significance ($P_{adjusted} = 0.452$). When comparing Group 1 and 2 combined with Group 3, there is a significant effect [$t(51) = 3.160$; $P = 0.001$].

In more detail, at the facet level Anxiety, Angry Hostility, and Depression as facets of Neuroticism show a significant *F*-value in the ANOVA (Table III) after Bonferroni correction. Post hoc tests showed significantly higher scores for Group 1 compared with Group 2 and Group 3 for Anxiety, Angry Hostility, and Depression.

Agreeableness barely missed a significant result in the ANOVA ($P_{adjusted} = 0.052$) after controlling for the inflated type I error. However, when comparing all unsuccessful candidates (Group 1 and 2 combined) with the successful candidates (Group 3), the difference was significant [$t(49) = 2.308$; $P = 0.021$]. The ANOVA of the remaining Big Five factors (Extraversion, Openness to Experience, and Conscientiousness) failed to reach significance as well.

The results of the correlation analyses with the TSS data are presented in **Table IV**. For the TSS, the scales Achievement and Vitality were significantly correlated with the OAR. All remaining correlations with TSS scales did not reach significance.

None of the Big Five factors correlated significantly to the OAR. Extraversion showed the highest correlation, but was not significant on a 5% α-level ($r = 0.132$; $P = 0.069$). At the facet level, only the facet Openness to Actions was significantly correlated with the OAR ($r = 0.142$; $P = 0.049$).

Table II. Means (M) and SDs on All TSS and NEO-PI-R Scales Separately for Groups 1, 2, and 3.

	GROUP 1*		GROUP 2†		GROUP 3‡	
SCALE	M	SD	M	SD	M	SD
Achievement	14.65	3.03	14.31	3.28	15.20	2.69
Instability	5.70	3.18	5.12	3.13	4.76	2.75
Rigidity	14.31	4.60	13.85	5.04	13.91	4.43
Extraversion	16.01	4.48	15.67	4.77	17.07	3.94
Aggressiveness	6.25	4.06	6.40	3.82	6.13	3.79
Vitality	14.91	4.41	14.01	4.46	15.30	3.82
Dominance	14.39	4.19	14.08	4.47	14.85	4.43
Empathy	13.51	3.14	13.49	3.25	13.22	3.41
Spoiltness	4.73	2.42	4.25	2.23	4.67	2.54
Mobility	9.09	2.88	9.03	2.82	9.59	2.41
Openness	6.90	4.07	7.42	4.42	7.15	3.69
Neuroticism	49.80	15.69	45.35	15.27	41.57	14.22
Extraversion	126.89	14.31	126.12	12.88	130.87	13.10
Openness to Experience	128.71	15.13	130.05	14.16	130.61	16.03
Agreeableness	132.85	14.24	134.50	13.49	138.09	15.11
Conscientiousness	140.96	13.90	140.03	15.44	142.20	13.71
Anxiety	8.53	3.72	7.44	3.74	6.52	3.53
Angry Hostility	6.89	3.64	5.90	3.18	5.30	3.37
Depression	7.10	3.77	6.27	3.56	5.22	3.15
Self-Consciousness	11.01	3.37	10.32	3.28	10.00	3.46
Impulsiveness	11.23	3.37	10.66	3.67	10.43	3.34
Vulnerability	5.04	2.90	4.76	2.81	4.09	2.72
Warmth	25.26	3.16	24.74	3.05	25.82	3.14
Gregariousness	20.40	3.90	19.86	3.91	20.80	3.19
Assertiveness	19.68	3.76	19.95	3.54	20.85	3.31
Activity	21.20	3.22	20.60	3.37	21.96	3.35
Excitement-Seeking	18.02	3.91	18.66	3.77	18.83	3.76
Positive Emotions	22.31	3.99	22.23	3.41	22.61	3.43
Openness to Fantasy	18.25	4.50	18.81	4.54	18.57	4.34
Openness to Aesthetics	20.62	4.79	20.40	4.61	20.35	5.44
Openness to Feelings	20.26	4.04	20.01	3.86	19.89	3.19
Openness to Actions	20.19	3.34	20.09	3.24	20.98	3.30
Openness to Ideas	26.06	3.66	26.99	3.35	26.78	3.66
Openness to Values	23.34	3.07	23.75	3.12	24.04	3.17
Trust	24.43	3.49	25.15	3.69	25.59	3.07
Straightforwardness	21.17	3.98	21.53	3.69	22.33	3.82
Altruism	25.34	3.02	25.30	3.17	25.96	2.74
Compliance	21.06	3.72	21.54	3.50	22.39	3.24
Modesty	18.84	4.30	18.80	4.02	19.02	4.86
Tender-Mindedness	22.01	3.32	22.17	2.97	22.80	3.02
Competence	24.49	2.98	25.23	3.10	24.93	2.57
Order	19.86	3.60	19.36	3.63	19.70	3.56
Dutifulness	26.61	3.02	26.83	2.99	27.54	2.35
Achievement-Striving	22.76	3.30	22.29	3.29	23.09	2.87
Self-Discipline	25.91	3.15	25.18	4.04	26.43	2.86
Deliberation	21.33	3.60	21.14	3.94	20.50	4.27

* $N = 710$; † $N = 146$; ‡ $N = 46$.

DISCUSSION

The main purpose of this study was to investigate personality differences between candidates in an astronaut selection who were successful and, therefore, recommended for training and those candidates who were unsuccessful in different phases of the selection procedure. The most notable effect in the data are the comparatively higher levels of Neuroticism or, more specifically, of the facets Anxiety, Angry Hostility, and Depression for candidates who failed phase 1 because they had deficiencies in at least one of the cognitive or psychomotor abilities that were tested or were not proficient enough in English. Completing a test battery that determines whether a candidate passes or fails is always prone to a degree of stress and anxiety. ESA astronaut selections happen in large intervals and there is practically only one chance in a lifetime to apply. This stress situation probably inhibited performance in some. Studies in the past have shown high relationships between Neuroticism and the level of test anxiety[2] and impairments of performance in stressful cognitive testing situations for highly neurotic individuals.[6] It is possible that Neuroticism has led to a certain test anxiety in some candidates who failed phase 1, thus reducing necessary capacities to perform well in this test situation. Future studies should investigate the level and impact of test anxiety in an astronaut selection.

Furthermore, the analyses did not include the individual cognitive and skill tests that were used in phase 1 of the selection and of which at least one must have been failed by those who did not pass that phase. This information may determine whether individual tests were affected specifically or all tests equally and whether the test situation itself was affected by Neuroticism, the preparation for the examination, or the acquisition of skills and knowledge in general. Thus, future analyses should also include data on tests for cognitive aptitudes and knowledge, as well as data on the preparation for tests in such a selection procedure.

Phase 2 was concerned with interpersonal communication, personality, and motivation. Thus, we expected stronger differences on personality scales related to interpersonal capacities (e.g., Agreeableness) between candidates who passed or failed phase 2 (Groups 2 and 3) as they were specifically screened for those capacities. This effect was not found in the data. Although the significant difference between successful and all unsuccessful candidates on Agreeableness indicates that somewhat more agreeable candidates were selected, they were not necessarily selected in phase 2, but throughout the whole selection procedure.

Table III. Results of ANOVAs for NEO-PI-R Factors and Facets Between Groups 1, 2, and 3.

	ANOVA		
SCALE	**F-VALUE**	**DF**	$P_{ADJUSTED}$
Neuroticism	19.98	2, 899	<0.001***
Agreeableness	6.60	2, 899	0.052
Anxiety	21.05	2, 899	<0.001***
Angry Hostility	15.74	2, 899	0.002**
Depression	15.68	2, 899	0.002**
Trust	8.79	2, 899	0.093

Only scales with $P_{adjusted} < 0.1$ are presented.
** $P_{adjusted} < 0.01$; *** $P_{adjusted} < 0.001$.

In summary, these results show some connections to the "Right Stuff" profile. As Musson et al.[22] reported, Neuroticism is negatively related to Mastery and Work Orientation and positively to Verbal Aggressiveness. Agreeableness is negatively related to Negative Instrumentality and Verbal Aggressiveness and positively related to Expressivity, all attributes associated with the "Right Stuff."

Neuroticism and Agreeableness are also among the sparse personality predictors for performance in confined environments[24] and overall astronaut effectiveness,[26] which underlines the validity and importance of these two personality traits in the actual work as an astronaut or in confined environments. However, other scales that were also found to be related to attributes of the "Right Stuff" (e.g., Conscientiousness) in previous studies[22] did not differ between successful and unsuccessful candidates in the present data.

In addition to investigating personality differences between the three candidate groups in the selection process, we also analyzed relationships between psychometric personality and the suitability rating. All candidates participating in phase 2 received an OAR at the end of the assessment which evaluates the overall suitability for the job given all the data that had been collected up to this point, including operational aptitudes that were tested in phase 1 as well as interpersonal capacities, social competency, and motivation that were assessed in phase 2.

Achievement and Vitality are the two TSS scales that are significantly correlated with the OAR in our analysis. While Achievement describes an ambitious, achievement-striving, focused, and industrious character, Vitality indicates physical robustness, strength, and athletic ambition.

Motivation is a big part of the "Right Stuff" and lies at the core of the definition[29] and can be distinguished in (general) achievement motivation and (specific) vocational motivation for the job at hand. While achievement motivation is often part of psychometric personality questionnaires, vocational motivation usually is not and has, therefore, to be assessed in an interview.

The relationship of Achievement with the OAR is plausible since the assessment of vocational and achievement motivation was a big part of the interview and high achievement motivation is a prerequisite for successful training and missions. Achievement (motivation) was also predictive for air traffic controller training success.[7] Additionally, training and missions are highly physically demanding and the significant correlation of Vitality with the OAR shows that Vitality was a source of variance in the evaluation of suitability of candidates despite the fact that the assessment of physical fitness is part of the medical selection afterwards.

Achievement and Vitality also resemble the trait of Instrumentality in the PCI with regards to the components of goal orientation, mastery, persistence, and engaging in no unhealthy behavior.[28] Instrumentality, as reported above, is part of the "Right Stuff" personality profile and was also found to be related to emotional stability during Antarctic missions.[28]

In contrast to previous research on psychometric personality in the selection of astronauts,[22] we found significant personality differences between successful and unsuccessful astronaut candidates and were able to show significant relationships of personality scales and an overall suitability rating. A possible explanation for these differences to the prior study could be the select-in approach that was used in this selection, which could have emphasized effects of general personality scales as opposed to the select-out focused approach used by NASA between 1990 and 1994.

However, the absolute magnitudes of the significant effects are, especially in view of the size of the sample, rather small and the selection decision was in large part not related to psychometric personality traits assessable with general personality questionnaires. This is not surprising considering the number of aspects that influence the OAR and the assumption

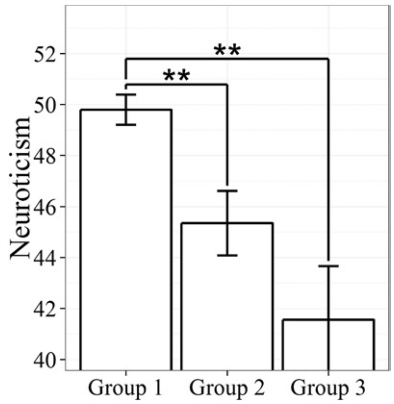

Fig. 1. Means and standard errors for Neuroticism (NEO-PI-R), separately for groups 1, 2, and 3; **$P < 0.01$.

Table IV. Pearson Correlation Coefficients of TSS Scales with Overall Assessment Rating (OAR).

SCALE	***r* OAR**
Achievement	0.188**
Instability	−0.033
Rigidity	0.020
Extraversion	0.076
Aggressiveness	−0.045
Vitality	0.176*
Dominance	0.127
Empathy	0.019
Spoiltness	0.093
Mobility	0.068
Openness	0.029

* $P < 0.05$; ** $P < 0.01$.

that correlations between personality traits and relevant behaviors cannot be expected to exceed 0.3.[19] The present sample is also a very specific sample with extreme personality characteristics and a limited range. The high degree of pre- and self-selection makes it difficult to find any large differences between successful and unsuccessful astronaut candidates.

Moreover, possible interactions between personality characteristics were also not considered in the analysis, but could definitely have had an impact on selection decisions and scoring. As indicated by the popularity of personality profiles or clusters like the "Right Stuff," "Wrong Stuff," and "No Stuff," not a high score on an isolated personality scale but the combination of high and low values on multiple personality scales can be beneficial.

The results lead to the conclusion that other factors like aspects of personality that are not assessed by personality inventories, but were relevant in the Assessment Center exercises or the interview must have influenced the OAR and the selection decision. Vocational motivation is one of these aspects and a very important one considering the dedication and persistence that is required in training and during missions.

As prior studies have suggested, personality traits were not predictive of selection success, but of overall performance.[9,22] Future studies need to further determine the relationship between personality traits and performance and need to indicate which psychometric personality traits should be focused more on in the selection of astronauts. The number of astronauts performing space missions for ESA is too small to compute reliable quantitative analyses and could potentially threaten the anonymity of the astronauts. Thus, other ways, like simulation studies, have to be used to perform validations of the applied instruments.

ACKNOWLEDGMENTS

Authors and affiliation: Justin M. Mittelstädt, M.Sc., Yvonne Pecena, Ph.D., Viktor Oubaid, Ph.D., and Peter Maschke, Ph.D., German Aerospace Center (DLR), Hamburg, Germany.

REFERENCES

1. Campbell JS, Castaneda M, Pulos S. Meta-analysis of personality assessments as predictors of military aviation training success. Int J Aviat Psychol. 2010; 20(1):92–109.
2. Chamorro-Premuzic T, Ahmetoglu G, Furnham A. Little more than personality: dispositional determinants of test anxiety (the Big Five, core self-evaluations, and self-assessed intelligence). Learn Individ Differ. 2008; 18(2):258–263.
3. Chidester TR, Helmreich RL, Gregorich SE, Geis CE. Pilot personality and crew coordination: Implication for training and selection. Int J Aviat Psychol. 1991; 1(1):25–44.
4. Costa PT, McCrae RR. Revised NEO Personality Inventory (NEO-PI-R) and Five Factor Inventory (NEO-FFI) Professional Manual. Odessa (FL): Psychological Assessment Resources; 1992.
5. DeYoung CG, Quilty LC, Peterson JB. Between facets and domains: 10 aspects of the Big Five. J Pers Soc Psychol. 2007; 93(5):880–896.
6. Dobson P. An investigation into the relationship between Neuroticism, Extraversion and cognitive test performance in selection. International Journal of Selection and Assessment. 2000; 8(3):99–109.
7. Eißfeldt H, Keye D, Conzelmann K, Grasshoff D, Pecena Y, Heintz A. Validity report: selection of ab-initio air traffic controllers for DFS Deutsche Flugsicherung GmbH (DLR Forschungsbericht, ISRN DLRFB-2013-02). Hamburg (Germany): German Aerospace Center (DLR); 2013.
8. Fitzgibbons A, Davis D, Schutte PC. Pilot personality profile using the NEO-PI-R. Hampton (VA): NASA; 2004. Report No.: NASA/TM-204-213237.
9. Grant I, Eriksen HR, Marquis P, Orre IJ, Palinkas LA, et al. Psychological selection of Antarctic personnel: the "SOAP" instrument. Aviat Space Environ Med. 2007; 78(8):793–800.
10. Helmreich RL, Spence JT, Wilhelm JA. A psychometric analysis of the personal attributes questionnaire. Sex Roles. 1981; 7(11):1097–1108.
11. Hörmann HJ, Maschke P. On the relation between personality and job performance of airline pilots. Int J Aviat Psychol. 1996; 6(2):171–178.
12. Judge TA, Rodell JB, Klinger RL, Simon LS, Crawford ER. Hierarchical representations of the five-factor model of personality in predicting job performance. J Appl Psychol. 2013; 98(6):875–925.
13. Kanas N. Group interactions during space missions. Aviat Space Environ Med. 2004; 75(7):C3–C5.
14. Kanas N. Humans in space: the psychological hurdles. Cham (Switzerland): Springer; 2015.
15. Kanas N, Manzey D. Space psychology and psychiatry, 2nd ed. Dordrecht (Holland): Springer; 2008.
16. Maschke P. Temperament Structure Scales (TSS). Oberpfaffenhofen (Germany): European Space Agency; 1987. Tech.Rep.No.: ESA-TT-1069.
17. Maschke P, Oubaid V, Pecena Y. How do astronaut candidate profiles differ from airline pilot profiles? Aviation Psychology and Applied Human Factors. 2011; 1(1):38–44.
18. McFadden TJ, Helmreich RL, Rose RM, Fogg LF. Predicting astronaut effectiveness: a multivariate approach. Aviat Space Environ Med. 1994; 65(10, Pt. 1):904–909.
19. Mischel W. Continuity and change in personality. Am Psychol. 1969; 24(11):1012–1018.
20. Mittelstädt JM, Pecena Y, Oubaid V, Maschke P. Construct validity of the Temperament Structure Scales within the Big Five framework in aerospace selection. Aviat Psychol Appl Hum Factors. 2016; in press.
21. Musson DM, Helmreich RL, Sandal GM. Baseline personality comparison between astronauts and Antarctic personnel: implications for generalization of psychological research findings. In: 55th International Astronautical Congress; 4-8 October, 2004; Vancouver, Canada. Reston (VA): AIAA; 2004.
22. Musson DM, Sandal GM, Helmreich RL. Personality characteristics and trait clusters in final stage astronaut selection. Aviat Space Environ Med. 2004; 75(4):342–349.
23. Palinkas LA, Gunderson E, Holland A, Miller C, Johnson J. Predictors of behavior and performance in extreme environments: the Antarctic Space Analogue Program. Aviat Space Environ Med. 2000; 71(6): 619–625.
24. Palinkas LA, Keeton KE, Shea C, Leveton LB. Psychosocial characteristics of optimum performance in isolated and confined environments. Hanover (MD): NASA; 2011. Tech. Rep. NASA/TM-2011-216149.
25. Pecena Y, Keye D, Conzelmann K, Grasshoff D, Maschke P, et al. Predictive validity of a selection procedure for air traffic controller trainees. Aviation Psychology and Applied Human Factors. 2013; 3(1):19–27.
26. Rose RM, Fogg LF, Helmreich RL, McFadden TJ. Psychological predictors of astronaut effectiveness. Aviat Space Environ Med. 1994; 65(10, Pt. 1): 910–915.
27. Sandal GM, Endresen IM, Vaernes R, Ursin H. Personality and coping strategies during submarine missions. Hum Perf Extrem Environ. 2003; 7(1):29–42.
28. Sandal GM, Vaernes R, Bergan T, Warncke M, Ursin H. Psychological reactions during polar expeditions and isolation in hyperbaric chambers. Aviat Space Environ Med. 1996; 67:227–34.
29. Santy PA. Choosing the Right Stuff: the psychological selection of astronauts and cosmonauts. Westport (CT): Praeger; 1994.