

Wright State University

CORE Scholar

---

International Symposium on Aviation  
Psychology - 2021

International Symposium on Aviation  
Psychology

---

5-1-2021

## The Role of a Group Assessment Centre in the Selection of Ab Initio Air Traffic Controllers

Jennifer Eaglestone

Alexander W J Arnold

Follow this and additional works at: [https://corescholar.libraries.wright.edu/isap\\_2021](https://corescholar.libraries.wright.edu/isap_2021)



Part of the [Other Psychiatry and Psychology Commons](#)

---

### Repository Citation

Eaglestone, J., & Arnold, A. W. (2021). The Role of a Group Assessment Centre in the Selection of Ab Initio Air Traffic Controllers. *38th International Symposium on Aviation Psychology*, 19-24.  
[https://corescholar.libraries.wright.edu/isap\\_2021/4](https://corescholar.libraries.wright.edu/isap_2021/4)

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2021 by an authorized administrator of CORE Scholar. For more information, please contact [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

## THE ROLE OF A GROUP ASSESSMENT CENTRE IN THE SELECTION OF AB INITIO AIR TRAFFIC CONTROLLERS

Jennifer Eaglestone

Netherlands Aerospace Centre (NLR) previously Air Traffic Control the Netherlands (LVNL)  
Amsterdam, The Netherlands

Alexander W. J. Arnold

Air Traffic Control the Netherlands (LVNL)  
Schiphol, The Netherlands

The selection of Air Traffic Controllers (ATCOs) is known to be extensive and highly selective. Air Traffic Control the Netherlands (LVNL) has a six-stage procedure for ab initio applicants. It includes two rounds of cognitive ability testing, personality testing, two job sample tests, an interview, and an assessment centre (AC). This study examines the construct- and predictive validity of the AC using 15 dimension-scores as predictor variables, with the final score on a high-fidelity job sample test as the dependent variable (criterion). A Principal Component Analysis (PCA) of the 15 dimension-scores resulted in two components, one (inter)personal behaviour component and one performance component. The performance component was found to have a stronger correlation with the cognitive ability tests than the (inter)personal behaviour component. Eight of the 15 AC predictor variables had a significant positive correlation with the high-fidelity job sample test.

Air Navigation Service Providers (ANSPs) are known to use extensive selection procedures before deciding which applicants can start training to become an Air Traffic Controller (ATCO). Training for ATCOs is long and expensive and, as with all such professions, any reduction in attrition rates means costs will be significantly reduced (Martinussen et al. 2000). Psychological assessment for Air Traffic Control (ATC) applicants has taken place since the early 60s, however up until the 90s it focussed mainly on cognitive attributes (Hättig, 1991). Job analyses have however underlined the importance of personality-related worker attributes in ATC with ATCO performance not only depending on technical knowledge and cognitive ability, but also on characteristics related to personality (Wium & Eaglestone, 2021).

Personality related factors are most often assessed using personality questionnaires and interviews. Damitz et al. (2003) however pointed out that these methods probably only assess personality related aspects of performance by evaluating signs of behaviour, as opposed to assessing actual behaviour. It assumes that how a candidate describes him or herself is predictive of the way that he or she will actually act. A broader approach, in which actual behaviour is also observed, would complement evaluation using a questionnaire and/or interview.

Assessment centres have been used since the 1950s to assess behaviour (Cascio and Silbey, 1979) and by the mid-70s assessment centres were, according to Mitchel (1975), being used widely within industry for the selection of managers. The use of assessment centers in selection in aviation is, however, comparatively new. In 1994 the German Aerospace Centre (DLR) developed one of the first assessment centres for pilot selection (Hörmann et al., 1997) and in 1996 an assessment centre was also added to the ATCO selection at DLR (Pecena, 2003).

In 2005 Air Traffic Control the Netherlands (LVNL) started using an assessment centre (AC) as part of their selection procedure for ab initio ATCOs. The AC in its current form has been used since 2014 and is the fourth stage of ATCO selection. In total candidates take part in six selection stages (hurdled approach), comprising of two rounds of cognitive ability testing, a low-fidelity job sample, the AC, an interview combined with a personality questionnaire and lastly a high-fidelity job sample test.

### **Aim and hypothesis**

This study aims to undertake a step in the validation (predictive and construct validity) of the assessment centre as a predictive selection tool for the selection of ab initio ATCOs at LVNL. Even though assessment centres can be valid predictors of job success (e.g., Hermelin et al., 2007) the construct validity of assessment centres has often been questioned (e.g., Sackett & Dreher, 1982). In 1990, Shore suggested that construct validity might be increased if single rating items were grouped into wider categories (Shore et al., 1990). He felt that assessors could probably only distinguish between interpersonal related behaviour and performance related behaviour and not between a vast number of competences or behaviours.

More recently Damitz et al. (2003) also stated that ratings of single behavioural dimensions have only low convergent and discriminant validity. While validating an assessment centre for pilot selection, Damitz et al. found that a Principal Component Analysis (PCA) clearly showed an interpersonal component and a performance related component in the assessment centre ratings. In this study we therefore expect the 15 Assessment Centre (AC) dimensions to consist of two components, an interpersonal and a performance related component (hypothesis 1a). Those AC dimensions in the performance (task-related) component should be more reliant on cognitive ability, whereas interpersonal behaviour is considered distinct from cognitive ability. Therefore, the AC dimensions in the performance component are expected to have a higher correlation with the cognitive aptitude tests than the AC dimensions in the interpersonal component (hypothesis 1b).

A lack of standardised training results that could be used as a criterion (due to multiple changes in training) fuelled the decision to use the LVNL high-fidelity job sample test (ACT) as the dependent measure when assessing predictive validity in this study. Not only does this job-sample have a high resemblance to actual ATCO training (taking place in the training simulator) but it also an expensive and time-consuming selection round, making it a necessity to only allow candidates with a high chance of passing to take part. The ACT job sample test is taken by applicants in the last stage of selection and candidates are taught to control traffic in an approach simulator for three days. After the three days they are scored by ATCO instructors on both behaviour and aptitude in the simulator. The 15 assessment centre dimension scores are expected to have a positive significant correlation with the final score on the ACT job sample test (hypothesis 2).

## **Method**

### **Sample**

This study was conducted using data from ab initio applicants for air traffic controller training at LVNL. The sample contained 1158 candidates who took part in the AC between October 2014 and July 2019. The average age is 23.4 years ( $SD = 3.4$ ) with a range from 17

years to 31 years. A total of 18.4% of the sample is female. Of these 1158 candidates, 160 took part in the ACT job sample test. These were candidates that had passed all previous selection stages including the assessment centre. Average age for this group is 24.2 ( $SD = 3.3$ ) years with a range from 17 years to 30 years (16.4% of the sample is female). ACT job sample data was collected between October 2016 and July 2019. This sample was used for the analysis of predictive validity.

### **Predictors and Criteria**

The LVNL assessment centre consists of three exercises: two group discussions (in groups of 4 candidates) and one individual exercise. Each of which is scored using the same 15 item score-form. All items are scored on a 4-point scale by a psychologist, whereby a maximum of 2 exercises per candidate are scored by the same psychologist. The three scores for each of the items were summed to create 15 dimension-scores (predictor variables). An overall score was also calculated for the assessment centre.

Performance on the LVNL high fidelity Job Sample test (ACT) is the dependent variable (criterion) in this study. Candidate performance during the ACT job sample was scored by ATCOs using 19 items (for example, planning, decisiveness, information processing) on a 6-point scale. The scores on the 19 items are summed to create the ACT final score. The dependent variable consisted of this final score.

In conjunction with Hypothesis 1b, scores from cognitive ability tests (round 2 of selection) were used. This selection round uses four subtests of the First European ATC Selection Test (FEAST) battery (Rathje, 2004) namely Heading and Range (SAHR), ATC Planning (SAP), Sorting Ability (SORT) and Visualisation (FOLD). These tests are administered with a cut-off per test and all subtests must be passed to be able to take part in the assessment centre.

### **Analyses**

In order to investigate the construct validity of the AC, a PCA was used with varimax rotation and a Kaiser normalization. The analysis was conducted using the 15 final dimension scores. Furthermore, the dimensions in each of the components resulting from the PCA were summed to create a competence score and then correlated with the raw FEAST subtest scores from the second round of selection. To assess the predictive validity, the 15 assessment dimension-scores were correlated with the ACT job sample final score. SPSS v25 was used for all analyses.

## **Results**

To assess hypothesis 1, a PCA with varimax rotation and Kaiser normalisation was used. A two factor solution was chosen based on an inspection of the scree plot and eigenvalues ( $> 1.0$ ) accounting for 49.9% of the variance. The results are presented in Table 1. Component 1 is composed of not only dimensions that would be categorised as interpersonal but seems to be somewhat broader and focussed on personal behaviour in general. Component 2 contains the more task-related dimensions and can indeed be classified as a performance component.

Processing speed however does not fit this pattern entirely. It scores relatively high on both components, with the highest loading on the (inter)personal component.

Table 1  
*PCA of Assessment Centre dimensions*

Assessment Centre dimensions	Component	
	1	2
Presence	.87	
Initiative	.85	
Assertive	.75	
Decisiveness	.68	
Flexibility	.48	
Accountability	.40	
Teamwork	.39	
Composure	.44	
Concentration	.45	.58
Argument effectiveness		.66
Processing speed	.55	.43
Solution effectiveness		.66
Prioritising		.67
Accuracy		.78
Structure		.75

*Note:* Factor loadings < 0.35 have been omitted for clarity.

To assess hypothesis 1b, two final assessment centre ratings were calculated for the two PCA components by summing the related dimension scores, creating an (inter)personal competence score and a performance competence score. These were correlated with the four cognitive ability subtests to analyse the relationship between cognitive ability and assessment centre performance. As the assessment centre comes after the cognitive ability tests in the selection procedure the sample was severely restricted in range. Table 2 shows the (uncorrected) correlations. For the SORT and FOLD subtests the correlation is significantly higher for the performance competence than for the (inter)personal competence.

Table 2  
*Correlations of assessment centre scores with the FEAST cognitive ability subtests*

	FEAST cognitive ability subtests			
	SAHR	SAP	SORT	FOLD
(Inter)personal behaviour competence (1)	.04	.02	.03	-.07*
Performance competence (2)	.08*	.08*	.13**	.08*

*Note:*  $n = 832$ . \* $p < .05$ , \*\* $p < .01$ . Difference  $r_{1, \text{subtest}}$  and  $r_{2, \text{subtest}}$  is sig. for SORT\* and FOLD\*\*

To study the predictive validity of the assessment centre (hypothesis 2) the overall AC rating and an (Inter)personal and Performance rating were correlated with the dependent variable – the ACT job sample final score. The three overall ratings all correlated significantly with the ACT final score. Furthermore eight of the 15 AC dimensions showed a significant correlation with the ACT job sample final score. Results are presented in table 3.

Table 3

*Correlations between Predictor and Criterion Measures (ACT job sample final score)*

	<i>r</i> (uncorr.)		<i>r</i> (uncorr.)
<i>Overall ratings</i>		<i>Dimension ratings</i>	
Overall Assessment centre rating	.33**	Presence	.16*
(Inter)personal	.17*	Initiative	.09
Performance	.32**	Assertive	.09
		Decisiveness	.08
		Teamwork	.08
		Flexibility	.07
		Accountability	.06
		Concentration	.27**
		Composure	.20**
		Argument effectiveness	.25**
		Processing speed	.09
		Solution effectiveness	.24**
		Prioritising	.29**
		Accuracy	.25**
		Structure	.20**

*Note: n = 155, \*p < .05, \*\*p < .01 (two-tailed test).*

### Discussion

The results of this study support earlier findings (Damitz et al. 2003, Shore et al. 1990) that assessors can distinguish between interpersonal behaviour and performance behaviour, although the interpersonal component here was somewhat broader and also encompassed worker-related personal behaviour such as composure, accountability and initiative that may not be considered *interpersonal*. The exception in this study was the dimension ‘processing speed’ which had a higher loading on the (inter)personal component than on the performance component, while one could argue that it is a more task/performance related dimension. This could be a sign of the difficulty with operationalisation of this dimension, making it difficult to evaluate. What assessors seem to be observing here is something more personality related. Correlations between the cognitive ability tests and the performance and (inter)personal behaviour dimensions have provided us with some evidence of construct validity.

Predictive validity was shown by the significant correlation between the assessment centre scores and the ACT final score. Although the separate (inter)personal behaviour dimensions do not correlate significantly, the combined score of the (inter)personal competence does show a significant correlation with the ACT final score. Restriction of range here is also greater than for the performance competence as the assessment centre is followed by an interview rating very similar (inter)personal dimensions.

Altogether it seems that the assessment centre is a promising tool in the selection of ab initio ATCOs at LVNL, however it is important that a full validation is carried out using training and job performance criteria.

## References

- Cascio, W. F., & Silbey, V. (1979). Utility of the assessment center as a selection device. *Journal of Applied Psychology, 64*(2), 107–118. doi: 10.1037/0021-9010.64.2.107
- Damitz, M., Manzey, D., Kleinmann, M. & Severin, K. (2003), Assessment Center for Pilot Selection: Construct and Criterion Validity and the Impact of Assessor Type. *Applied Psychology, 52*, 193-212. doi: 10.1111/1464-0597.00131
- Hättig, H. J. (1991). Selection of air traffic control cadets. In R. Galand, and A. D. Mangelsdorff (Eds.), *Handbook of Military Psychology* (p. 115-129). New York, NY: John Wiley & Sons.
- Hermelin, E., Lievens, F. & Robertson, I. T. (2007). The validity of assessment centres for the prediction of supervisory performance ratings: A meta-analysis. *International Journal of Selection and Assessment, 15*, 405-411. doi: 10.1111/j.1468-2389.2007.00399.x
- Hörmann, H.-J., Manzey, D., Maschke, P., & Pecena, Y. (1997). Behavior-oriented assessment of interpersonal skills in pilot selection. In R.S. Jensen (Ed.), *Proceedings of the ninth international symposium on aviation psychology: Vol. 2* (pp. 1110-1115). Columbus, OH: The Ohio State University.
- Martinussen, M., Jenssen, M., & Joner, A. (2000). Selection of air traffic controllers: Some preliminary findings from a meta- analysis of validation studies. Paper presented at the 24th EAAP conference (European Association for Aviation Psychology), Crieff, Scotland.
- Mitchel, J. O. (1975). Assessment center validity: A longitudinal study. *Journal of Applied Psychology, 60*(5), 573–579. doi: 10.1037/0021-9010.60.5.573
- Pecena, Y. (2003). An assessment-center approach to ATCO selection – An evaluation study. In Proceedings of the second EUROCONTROL Selection Seminar (pp. 201-208). Luxembourg, 16.-18.Nov. 2003.
- Rathje, H., Golany, Z., and Eißfeldt, H. (2004). Pan-European Selection Test Battery for Air Traffic Control Applicants. In K. Goeters (Ed.), *Aviation Psychology: Practice and Research* (p. 171 – 201). Aldershot: Ashgate
- Sackett, P. R., & Dreher, G. F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology, 67*(4), 401-410. doi: 10.1037/0021-9010.67.4.401
- Shore, T. H., Thornton, G. C., & McFarlane Shore, L. M. (1990). Construct validity of two assessment center dimension ratings. *Personnel Psychology, 43*, 101-116.
- Wium, J., & Eaglestone, J. (2021). Job analysis for air traffic control: A review of the literature from 1949-2020. Unpublished manuscript.