



**HAL**  
open science

# **Predictive validity of cognitive abilities for air traffic controllers selection among proficient scientific candidates**

Damien Mouratille, Franck Amadieu, Nadine Matton

## ► **To cite this version:**

Damien Mouratille, Franck Amadieu, Nadine Matton. Predictive validity of cognitive abilities for air traffic controllers selection among proficient scientific candidates. *Applied Ergonomics*, 2026, 131, pp.104666. <10.1016/j.apergo.2025.104666>. <hal-05327331>

**HAL Id: hal-05327331**

**<https://hal.science/hal-05327331v1>**

Submitted on 28 Oct 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

**Predictive Validity of Cognitive Abilities for Air Traffic Controllers Selection  
among Proficient Scientific Candidates**

Damien Mouratille<sup>1</sup>, Franck Amadiou<sup>2</sup>, and Nadine Matton<sup>1,2</sup>

<sup>1</sup> Fédération ENAC ISAE-SUPAERO ONERA

Université de Toulouse

France

<sup>2</sup> Laboratoire CLLE (CNRS 5263)

Université de Toulouse

France

<sup>3</sup> Author Note

Correspondence concerning this article should be addressed to Damien Mouratille,  
ENAC, 7 avenue Edouard Belin, 31055, Toulouse, France. E-mail :

[damien.mouratille@enac.fr](mailto:damien.mouratille@enac.fr)

**Abstract**

Most air traffic controllers (ATCOs) are selected after high school graduation, partly based on cognitive abilities. On the contrary, French ATCOs are selected after an intensive scientific post-secondary program without undergoing cognitive assessment. This study examines the potential incremental validity of cognitive predictors in ATCO training for such proficient applicants. Initial and unit training results were available for 414 and 277 ATCO students, respectively. All students took cognitive ability tests, encompassing visuospatial processing, quantitative knowledge, and work samples. We compared a baseline regression model with academic performance as sole predictor to an extended model with the addition of the cognitive abilities composite score. The  $R^2$  of the extended model increased significantly, by 0.04 ( $p < .001$ ) and 0.06 ( $p < .001$ ), respectively, for predicting practical initial and unit training outcome. In conclusion, even for scientific proficient candidates, cognitive assessment can enhance the success rate of ATCO training.

## **Predictive validity of cognitive abilities for Air Traffic Controllers Selection among proficient scientific candidates**

### **Introduction**

In safety-critical professions, such as those of air traffic controllers (ATCOs), candidates' selection undergoes a rigorous process that is closely examined (Eaglestone et al., 2022). Indeed, ATCO training is costly, complex, time-consuming, and critical as it is essential for ensuring the safe and efficient operation of the air traffic management system. Meanwhile, the demand for ATCOs is projected to rise alongside the anticipated growth in air traffic between now and 2040 (FAA, 2021). Currently, this demand is already manifesting itself in a mismatch between current and required staffing levels, resulting in air traffic disruptions such as flight delays or re-routings. For example, staffing problems caused 26% of European flight delays in 2021, with French en-route ATCO staffing shortages accounting for over one-third of this delay (Eurocontrol, 2022). A comparable situation is observed in the United States, where more than 90 percent of the US ATCOs facilities operate below the Federal Aviation Administration's recommended staffing levels (Krolik, 2025). Furthermore, a specific concern in ATCO training is attrition, which refers to students resigning or being stopped during their training. This results in financial losses, organizational issues, and potential psychological challenges for the students. Attrition is driven by multiple factors, including declining motivation, misalignment between the profession and student interests, and performance failures, among others. These are the reasons behind the long-standing tradition of selection in aeronautics (DuBois, 1947). Meta-analyses conducted on ATCOs' selection revealed significant predictive validities for various cognitive ability predictors (Martinussen et al., 2000; Mouratille et al., 2022). However, most ATCO predictive validity studies have been conducted on candidates who have recently graduated from high school (e.g., Mouratille et al., 2022). This study aimed at testing whether cognitive abilities would predict ATCO training success in a population of scientific proficient students. Indeed, one could argue that such students may already exhibit high cognitive abilities, potentially limiting the incremental validity of these measures.

In the general population, cognitive abilities are known to be linked to job performance (Lang & Kell, 2020; Nye et al., 2022; Ree & Carretta, 2022; Sackett et al., 2024; Salgado & Moscoso, 2019). Salgado & Moscoso (2019) updated the famous meta-analysis by Hunter and Hunter (1984) and revealed similar observed correlation values ( $r = .31$  for moderate complexity jobs) between scores on general cognitive abilities and professional success criteria from 1950 to 1985. Furthermore, it is known that for complex tasks requiring the processing of new information, like during training of complex cognitive skills, broad cognitive abilities have the best predictive validity (Ackerman & Beier, 2018). A meta-analysis

including more than 100,000 participants highlighted that cognitive ability is a substantial prerequisite for scholastic success measured through course grades (Roth et al., 2015). In U.S. graduate school environments, Woo et al. (2023) argued in favor of cognitive predictors instead of academic variables such as Grade Point Average. In the same way, cognitive abilities were found to have incremental validity over prior educational attainment in British medical schools (Paton et al., 2022) and German psychology schools (Watrin et al., 2022).

Predictors of a selection process should be chosen according to content and predictive validity. Content validity relies on job or task analysis, ensuring that the abilities evaluated during selection are relevant to the job. A recent systematic review of ATCO job analyses (Wium & Eaglestone, 2022) highlighted diverse methodologies and taxonomies used across air navigation service providers and working conditions. However, job activities and requirements remain consistent, as ATCOs are responsible for monitoring and regulating air traffic to ensure safe and efficient aircraft movement. Furthermore, ATCO is considered among the occupations that require the highest levels of cognitive abilities (Wolfram, 2023). From a cognitive point-of-view, the central aspect of this occupation is the dynamic process of acquiring, processing, and sharing information from multiple sources (Goeters et al., 2004). This process involves concurrent multitasking with potential adaptations based on the integration of new data. Goeters et al. (2004) highlighted the critical importance of divided attention, selective attention, monitoring, listening comprehension, vigilance, and perceptual speed. In the context of US ATCOs, visuo-spatial processing, fluid reasoning, working memory, speed processing, and long-term memory have been regarded as the most important cognitive abilities (Broach, 2013). As a result, ATCO trainees are usually selected on their cognitive abilities. For example, the FEAST battery, provided by Eurocontrol and currently used by 56 organizations in the world, comprises numerous cognitive ability tests (FEAST, 2024).

Predictive validity of cognitive ability test scores has been highlighted through meta-analyses of ATCO selection. By combining 35 independent studies, Martinussen et al. (2000) highlighted low-to-medium mean correlations ( $r \in [.10, .30]$ ) between training success and various predictors such as occupational knowledge, mathematical ability, multiple task performance, verbal ability, general ability, spatial ability, work sample (low-fidelity simulation of real job tasks) and composite scores (weighted aggregate of multiple test scores). More recently, Mouratille et al. (2022) aggregated 51 studies published between 1961 and 2021 ( $N = 65,839$  participants) and identified four broad cognitive abilities that showed large effects in the prediction of success in training ( $\rho > .30$ ): quantitative knowledge, processing speed, working memory, and visuo-spatial processing. Two other cognitive predictors showed large effects, namely work sample and cognitive composite. Moreover, predictive validity was strongly moderated by criterion nature. In general, the initial training is more precisely predicted than the unit training by

cognitive predictors ( $\rho = .52$  vs  $\rho = .20$ ). In the initial training, theoretical and practical grades are often averaged, although some studies specified the values for each type of assessment. For example, Pecena et al. (2013) reported small correlation values between cognitive abilities and theoretical grades ( $r \in [.10, .18]$ ) whereas Carretta and King (2008) found large correlation values ( $r \in [.20, .44]$ ). Interestingly, level of education exhibited a medium correlation with training success, but the highest level of education required was high school (Mouratille et al., 2022). In other words, most ATCO predictive validity studies were conducted on candidates who have recently graduated from high school. Thus, the predictive validity of cognitive abilities for candidates pre-selected on their scientific proficiency during post-secondary education remains an open question. Significant predictive validities of cognitive ability test scores have been found among a restricted population characterized by high cognitive abilities (Brown et al., 2021) or in high-stakes settings (Sackett et al., 2008). This link is highly stable from adolescence to late adulthood (Breit et al., 2024). In addition, it is very unlikely to find students with very low levels of cognitive abilities among those proficient in science (Zhou et al., 2024a), suggesting that such proficiency is associated with high cognitive ability.

Contrary to the majority of ATCO selections (Eaglestone et al., 2022; FEAST, 2024), the French ATCO selection process solely relies on scientific knowledge assessments and does not encompass any cognitive assessment. ATCO candidates are selected after a specific French degree, a two-year undergraduate degree, called "preparatory classes of higher engineering schools". This post-secondary education focuses on the learning of advanced scientific knowledge and skills in mathematics, physics, chemistry, and engineering science. It's a highly selective and intensive two-year program designed to prepare students for entrance exams to prestigious higher education institutions known as "grandes écoles". For information, in France, 91% of high school students graduate, and only 17% of them gain admission to preparatory classes (Ministry of Higher Education, Research and Innovation, 2020). Following the first year of preparatory class, success rate is approximately 75% as exemplified by a 72% rate in 2011 (Ministry of Higher Education, Research and Innovation, 2013). Moreover, approximately 10% of applicants for the ATCO entrance exam are selected for ATCO training, based on assessments in subjects such as mathematics and physics, among others. Hence, French ATCO applicants demonstrate a high level of proficiency in scientific knowledge and academic skills. Each year, around one hundred ATCO students are recruited with a European Class 3 Medical Certification. Throughout their ATCO training, students undergo assessments in both theoretical knowledge and practical skills. Practical assessments are carried out during simulation and are based on a standardized list of competencies requiring some level of proficiency. Among them, approximately 10% fail during training, mainly due to performance failure. This rate of 10% can be regarded as low in comparison to other ATCO selections (between 20% and 50%,

[Pecena et al., 2019](#); [Langr et al., 2019](#)). However, these failures lead to potential psychological problems for students (failure), organizational issues (one ATCO missing in operational center) and financial losses for the government and taxpayers (for example, training one US ATCO student costs around 250,000 \$ for two and a half years; [FAA, 2014](#)). They could be avoided with cognitive screening because some critical cognitive abilities may not be assessed by scientific knowledge assessments. In other words, the current selection process primarily evaluates acquired knowledge, particularly the broad ability of domain-specific knowledge. According to the Cattell-Horn-Carroll (CHC) theory, this ability is defined as 'specialized declarative and procedural knowledge' ([Schneider & McGrew, 2018](#)). However, evidence suggests this ability weakly predicts ATCO training success within high school populations ( $\rho = .19$ ; [Mouratille et al., 2022](#)). In France, advanced scientific proficiency, assessed after the two-year undergraduate degree in science, is considered to be a better predictor than scientific proficiency after high school because advanced scientific proficiency would be underpinned by higher cognitive abilities. For instance, physical scientists are among the occupations that demand the highest levels of cognitive abilities ([Wolfram, 2023](#)). Nevertheless, declarative knowledge alone might not fully capture the cognitive demands of ATCO training and work (real-time parallel processing, fast decision-making in unpredictable environments and managing high-stakes operational environments) even for proficient scientific candidates. Although domain-specific knowledge constitutes a foundational prerequisite, training success likely requires additional abilities, such as visuospatial abilities, working memory, attentional control, and stress tolerance ([Wium & Eaglestone, 2022](#)). Notwithstanding the ongoing debates ([Sackett et al., 2022](#); [Woods & Patterson, 2023](#)), the assessment of cognitive abilities is regarded as one of the most effective methods for increasing predictive validity in a selection process ([Kulikowski et al., 2025](#)).

The main purpose of this study was to investigate the incremental validity of cognitive predictors for ATCO training within a sample of French ATCO students who have been selected based on the results of a knowledge-based (mainly science) entrance exam. This is the first published longitudinal study focusing on this specific population. Performance at the ATCO entrance exam has been collected for each applicant. Moreover, at the beginning of initial training, all the selected ATCO students had to perform cognitive ability tests assessing four broad cognitive abilities highlighted by job analysis and meta-analysis: visuo-spatial processing, quantitative knowledge, working memory and processing speed. Finally, data related to ATCO training outcomes has been gathered upon completion of the initial training, encompassing grades on theoretical and practical assessments. For most of the participants, unit training outcomes were also available. Criterion-related validity evidence will be employed to test three hypotheses:

- 1) Hypothesis 1 (H1): Compared to the current selection process (i.e., academic performance), cognitive predictors will have incremental validity for predicting the theoretical grades during initial training.
- 2) Hypothesis 2 (H2): Compared to the current selection process (i.e., academic performance), cognitive predictors will have incremental validity for predicting the practical grades during initial training.
- 3) Hypothesis 3 (H3): Compared to the current selection process (i.e., academic performance), cognitive predictors will have incremental validity for predicting the pass/fail unit training outcome.

## Method

### Participants

The sample consisted of data from 414 ATCO trainees (age  $M = 21.5$  years,  $SD = 2.0$ ; 69% male; 31% female) selected between 2015 and 2021. These ATCO trainees were students at the French National School of Civil Aviation, the only ATCO school in France. Cognitive abilities were assessed at the beginning of initial training (after being selected). Since this was a low-stakes assessment, test-taking engagement indicator was computed based on response time and used in order to discard test-takers exhibiting disengaged behaviors (Rios et al., 2014). Low score from disengaged behavior would consequentially be an underestimation of true ability and should be avoided (data of two participants were discarded). Training outcome data was collected for all trainees during initial training, and was available, at the time of data collection (*i.e.* at the beginning of 2025), for 277 of them concerning unit training. For the remaining, the unit training is still underway and should be completed within two years. This study was conducted in accordance with the principles expressed in the Declaration of Helsinki.

### Material

Predictors included academic performance and cognitive abilities. The academic performance is based on the ATCO entrance exam. The results of the written tests (mathematics, physics, English, French, a supplementary test on a variable topic and a bonus optional test; these last two tests differ each year and can be on aeronautical knowledge or programming) and oral tests (mathematics, physics, English, French) were averaged and weighted according to the official coefficients (mathematics: 2, physics: 2, English: 3,

French: 2, compulsory option: 3, all oral tests: 2). This weighted average was then standardized by year of recruitment and corresponded to the ‘academic performance’ variable.

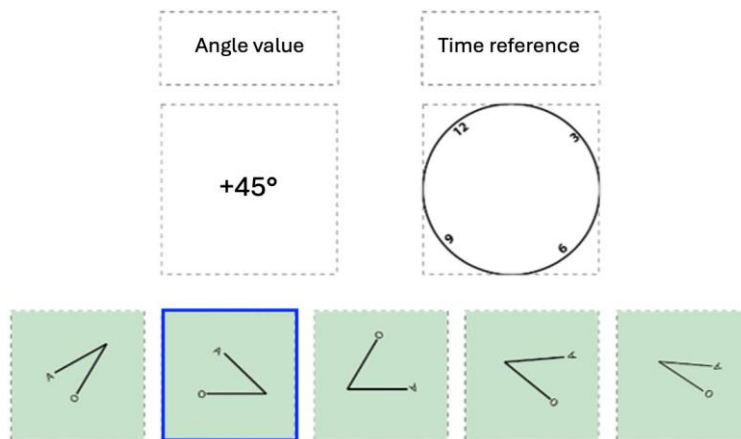
### *Cognitive ability tests*

Cognitive abilities were categorized based on CHC theory (Schneider & McGrew, 2018). Participants took four tests: a visuo-spatial processing test ( $Gv$ ), a quantitative knowledge test ( $Gq$ ), a pilot multitasking work sample test, and a low-fidelity ATC simulation which can also be considered as a multitasking work sample test. Both work sample tests required mainly two broad cognitive abilities, namely working memory ( $Gwm$ ) and processing speed ( $Gs$ ). Each of the scores were converted into a z-score by year of recruitment. In addition, one composite score was created by averaging all z-scores.

Visuo-spatial processing ( $Gv$ , the narrow ability being speeded rotation  $SR$ ) was assessed by a static test. Participants had to identify the angle corresponding to the time reference and the angle value indicated (see Figure 1). The number of correct items was counted for the 30 questions. Reliability of this test was computed and can be considered as good (Cronbach’s alpha was 0.83).

**Figure 1**

*Visuo-spatial test.*



*Note. The participants were presented with an angle value (either positive or negative) and a direction (clockwise or anticlockwise). The angle was always read from 0 to A. The objective was to identify the corresponding angle from the provided choices at the bottom. The correct response is the proposition in a blue box.*

Quantitative knowledge ability ( $Gq$ , the narrow ability being mathematical achievement  $A3$ ) was assessed by an arithmetic test. Participants had to solve mathematical operations. The number of correct items was counted for the 40 questions. Reliability of this test was computed and can be considered as good (Cronbach's alpha was 0.84).

Two multitasking work sample tests, or low-fidelity simulations, were used to assess working memory ( $Gwm$ ), related to controlled attention. Given the dynamic aspect of the tasks, both work sample tests also partly assessed processing speed ( $Gs$ ). During the ATC work sample (see Figure 2), participants had to guide two airplanes moving dynamically and simultaneously into two mazes by giving them instructions (relative headings) in the LABY Microworld (Imbert et al., 2014). Thus, participants had to divide their attention to deal with both aircraft at the same time. Airplanes should follow a defined trajectory with obstacle avoidance. Performance was calculated on the basis of the time spent outside this defined trajectory. This delay was summed up and reversed. Thus, a higher score corresponded to a better performance like all other predictors. Reliability of this test was computed and can be considered as questionable (split-half reliability value of 0.67 with a Spearman-Brown Correction).

**Figure 2**

*ATC multitasking work sample test.*



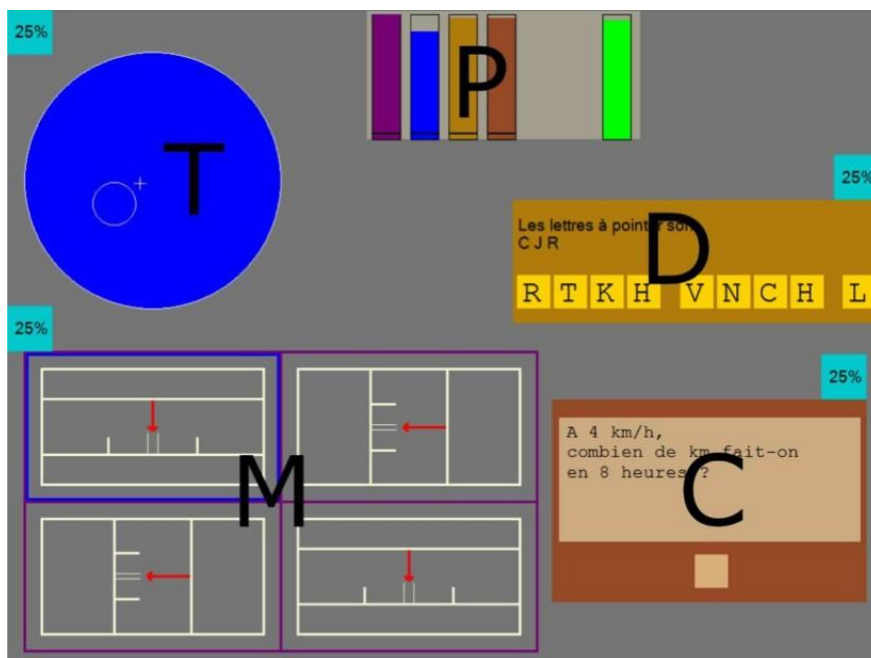
*Note. In this example, the participant was required to give instructions to two aircraft moving simultaneously (in red), using the menu on the bottom right (-20 is equal to a left turn of 20 degrees, +20 to a right turn of 20 degrees). Aircraft begin their flight on the green line and must end in the blue rectangle.*

The pilot work sample, labelled the Priority Management Task, was organized in six successive stages (Matton et al., 2016). Participants had to complete simultaneously four concurrent subtasks that were successively added (from only one subtask at the first stage to four subtasks at the fourth stage):

Monitoring, Tracking, Detecting and Calculating (see Figure 3). These tasks are cognitively similar to those performed in an aircraft. The ‘Monitoring task’ consisted in maintaining the levels of four gauges within an interval by using a first joystick. The ‘Tracking task’ consisted in keeping a cross positioned in a moving circle in an area bordered by a large circle, through the second joystick. The ‘Detecting task’ consisted in detecting the presence of three target letters in a block of nine letters. The ‘Calculating task’ consisted of simple arithmetic problems (e.g., deducing a distance from given speed and time). At each stage, instructions indicated that each subtask was equally important, and the number of concurrent subtasks increased from stage 1 (Monitoring subtask only) to stage 4 (all four subtasks concurrently). Two final stages varied the assigned priorities of the four subtasks. All the characteristics of the subtasks were the same for each participant. Each subtask required an action from the participant every 15 s at the same time. A score was computed on the basis of the mean performance of stage 3, 4, 5, and 6. Reliability of this test was computed and can be considered as good (split-half reliability value of 0.83 after a Spearman-Brown Correction).

**Figure 3**

*Pilot multitasking work sample (Priority Management Task).*



*Note. In Stage 1, participants were required to monitor four gauges (Task M). In Stage 2, they were tasked with simultaneously completing Tasks M and T, task T being a tracking task. In Stage 3, Task D, which involved detecting letters, was introduced. In Stage 4, participants were required to solve simple arithmetic problems (Task C) while completing the previous three tasks. During Stage 5 and 6, priorities of each task*

were modified (in this figure, each task has the same priority, 25%). Throughout the task, the participants were presented with an instantaneous performance display (P).

### Training outcomes

Data of the training outcome of all these ATCO students were collected after completion of initial training and unit training for the majority of them. Three different criteria were collected, two of which relate to the initial training (theoretical and practical part) and are continuous, and one relates to the unit training and is dichotomous. In our sample, all training failures were due to performance-related issues. Voluntary withdrawals were extremely rare and did not impact the analyses, as the few concerned participants were excluded from the dataset.

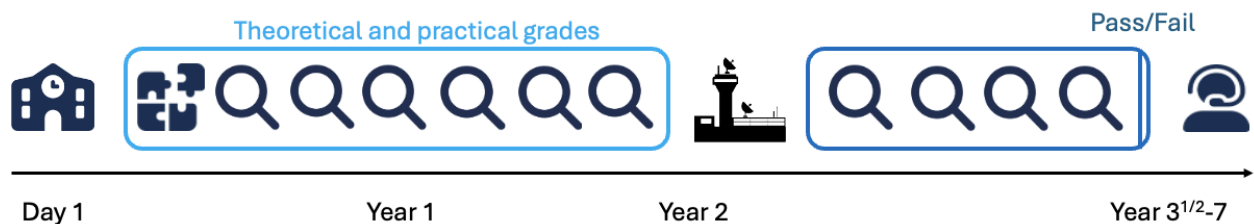
For the theoretical criteria, all initial training theoretical and practical exams have been averaged and z-scored by each class of students. Therefore, potential differences in exam difficulty across years were considered. At the end of the unit training, one final on-the-job exam is carried out and if the trainee succeeds, he becomes a certified ATCO. This criterion has the value ‘Success’ if the trainee successfully completed the unit training or the value ‘Failure’ if the trainee was required to stop or repeat the unit training.

### Procedure

The assessment was carried out in computer rooms during the first month of each class. For three hours, participants carried out the four tests. At the beginning of the testing sessions, participants were provided instructions, and their consent was obtained. Training outcomes were provided by the Air Traffic Management department of the school.

**Figure 4**

*Timeline of this longitudinal study.*



*Note. In Day 1, French ATCO students begin their initial training (after a selection process based on scientific knowledge assessments). During the first month of training, psychotechnical testing is carried out. Then, theoretical and practical competencies are assessed throughout the following two years. At the end of*

*Year 2, ATCO students are assigned to their operational center, after which unit training begins. During this unit training, competencies are continually assessed but the content, the frequency and the number are subject to variation depending on the center. After completing the unit training (its duration ranges from 1 to 5 years, depending on when the instruction team deems the student ready), ATCO students graduate and become “First Controller” (Certified Professional Controller).*

## Statistical analyses

Firstly, descriptive statistics were computed. Pearson correlations were computed for both continuous outcomes and point biserial correlations for dichotomous variables. Correlation values were categorized using the following rule:  $r > .28$  = large,  $r > .15$  = medium,  $r > .07$  = small (Bosco et al., 2015).

Secondly, multiple linear regressions were modelled for both continuous criteria and a two-stage hierarchical logistic regression was modelled to predict the unit training outcome<sup>1</sup>. In the first step, the variable ‘Academic performance’ was introduced. In the second step, two models were tested: either the introduction of the composite score (Model 2a), or the introduction of all individual variables (Model 2b). A chi-squared goodness-of-fit test was applied after each step to check the incremental validity of the step and  $R^2$  (or McFadden’s  $R^2$  for logistic regression) was used to measure the explained variance. The best model was chosen based on the Akaike information criterion (AIC). The change in total explained variance allowed to test the three hypotheses (Cohen, 1988). This change can be high ( $\Delta R^2 > .25$ ), moderate ( $\Delta R^2 > .09$ ), low ( $\Delta R^2 > .01$ ) or non-existent ( $\Delta R^2 < .01$ ).

Thirdly, k-fold cross-validation ( $k = 5$ ) was applied to the regression models to test the model generalization (de Rooij & Weeda, 2020). *Cv* package with root mean squared error (RMSE) metric and *boot* package with area under the curve (AUC) metric were used for linear and logistic regressions, respectively.

In addition, the confusion matrix was calculated with a threshold selected to maximize sensitivity given a minimal value of specificity using the *cutpointr* package.

## Results

Initial examination of the data concerning statistical assumptions (normality, multicollinearity, and homoscedasticity) was performed before applying the analyses and no violation was observed (all Variance Inflation Factors were  $< 1.2$ ). Extreme outliers were also checked with the *car* package in R. The

---

<sup>1</sup> As the classes were unbalanced, the Firth method was tested (Wang, 2014) but no difference was measured.

outlierTest() function identified two outliers in our dataset. However, these outliers correspond to instances of “failure” in our unit training outcome variable. Given the low frequency of failures, we chose to retain these data points as our objective is to understand factors associated to failure.

Table 1 shows the correlations between all variables. Academic performance correlated significantly only with theoretical grades in initial training, but not with practical grades or with the unit training outcome. All cognitive variables correlated significantly with initial training grades (except for quantitative knowledge and theoretical grades). All cognitive variables correlated significantly with unit training outcome. Effect size ranged from small to medium. Of the 412 students, 370 successfully completed the initial training (90%) and 42 failed (10%). Of the 370 students who completed the initial training, 256 successfully completed the unit training (92%), 21 failed (8%) and 93 were still in unit training. Control analyses were done to compare participants who finished the unit training and those who did not, and no significant difference was observed on our predictors. All the analyses undertaken in this study were also performed with gender as a control variable, with interaction effects and without extreme outliers, however it did not provide any substantial additional information. For the sake of conciseness, analyses with gender variable, interaction effects and without extreme outliers were not reported.

**Table 1**

*Descriptive Statistics and Intercorrelations*

Variables	N	M	SD	1	2	3	4	5	6	7	8
1. Academic performance	412	0.02	1.01	—							
2. Visuo-spatial processing	412	11.76	5.33	.05	—						
3. Quantitative knowledge	412	16.65	6.44	.09	.51***	—					
4. ATC work sample	412	163.70	69.57	-.01	.26***	.31***	—				
5. Pilot work sample	412	58.58	9.87	.07	.49***	.41***	.35***	—			
6. Composite score	412	-0.01	0.73	.07	.77***	.76***	.65***	.77***	—		
7. Initial training – Theoretical grades	397	0.04	0.95	.14**	.14**	-.01	.12*	.12*	.13*	—	
8. Initial training - Practical grades	397	0.05	0.93	.08	.14**	.10*	.16**	.19***	.20***	.59***	—
9. Unit training outcome	277	—	—	-.02	.12*	.15*	.18**	.23***	.23***	.22***	.25***

*Note.* N: number of participants, M: mean, SD: standard deviation. Pearson correlations except for correlations with unit training outcome (point-biserial). Two-tailed tests. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

### Initial Training - Theoretical grades

Table 2 shows the results of the linear regression analyses conducted to predict the theoretical grades during the initial training. In Model 1, academic performance was associated with higher theoretical grades ( $\beta = 0.15$ ,  $p = .002$ ) and was explaining 2% of the variance. In Model 2a, composite score was added to the model and accounted for 2% of the variance in theoretical grades above and beyond academic performance ( $p < .01$ ). Composite score was associated with the criterion ( $\beta = 0.14$ ,  $p = .005$ ). Therefore, in line with our expectations, cognitive predictors added incremental predictive validity for theoretical grades compared to academic performance alone (**H1**). In Model 2b, all cognitive abilities scores were introduced and accounted for 4% of the variance above and beyond academic performance ( $p < .001$ ). Visuo-spatial processing ( $\beta = 0.16$ ,  $p < .01$ ) was significantly associated with higher theoretical grades. Unexpectedly, quantitative knowledge was also associated but in the opposite direction ( $\beta = -0.12$ ,  $p = .04$ ). By removing only one of the 397 participants (who had very low theoretical grades and a very high quantitative knowledge score), this surprising result disappeared, and the association was no more significant ( $\beta = -0.11$ ,  $p = .05$ ). The best model in terms of AIC was Model 2b, the one with all individual scores of cognitive abilities (1039 < 1042) in which academic performance and visuo-spatial processing were significant predictors of theoretical grades. For the three models, RMSE and cross-validation RMSE were equivalent, generalizability of the three models can be expected.

**Table 2.**

*Linear regression models for theoretical grades during initial training.*

Predictors	Model 1		Model 2a		Model 2b	
	$\beta$	$p$	$\beta$	$p$	$\beta$	$p$
1. Academic performance	0.15 [0.06 – 0.25]	.002	0.14 [0.05 – 0.24]	.003	0.15 [0.05 – 0.25]	<.01
2. Composite Score			0.14 [0.04 – 0.24]	.005		
3. Visuo-spatial processing					0.16 [0.04 – 0.28]	<.01
4. Quantitative knowledge					-0.12 [-0.24 – -0.01]	.04
5. ATC Work Sample					0.08 [-0.03 – 0.18]	.15
6. Pilot Work Sample					0.07 [-0.04 – 0.19]	.23
$R^2$	.02		.04		.07	
$\Delta R^2$			.02 ( $p < .01$ )		.04 ( $p < .001$ )	

AIC	1048	1042	1039
RMSE	0.90	0.90	0.89
Cross-validation RMSE	0.91	0.90	0.90

### Initial Training - Practical grades

Table 3 shows the linear regression analyses for predicting the practical grades during the initial training. In Model 1, explaining 1% of the variance, academic performance was not associated with higher practical grades. In Model 2a, composite score was added to the model and accounted for 4% of the variance in practical grades above and beyond academic performance ( $p < .001$ ). In addition, the relationship between the composite score and the criterion was significant ( $\beta = 0.21$ ,  $p < .0001$ ). Thus, evidence is found to support **H2**. In Model 2b, all cognitive abilities scores were introduced and accounted for 5% of incremental variance above and beyond academic performance ( $p < .001$ ). The pilot multitasking work sample test was significantly associated with higher practical grades ( $\beta = 0.13$ ,  $p = .03$ ). However, the best model in terms of AIC was Model 2a, the one comprising only the composite score ( $1053 < 1058$ ). For the three models, RMSE and cross-validation RMSE were equivalent (or even better for Model 2b), generalizability of the models can be expected

**Table 3.**

*Linear regression models for practical grades during initial training.*

Predictors	Model 1		Model 2a		Model 2b	
	$\beta$	$p$	$\beta$	$p$	$\beta$	$p$
1. Academic performance	0.10 [-0.00 – 0.19]	.06	0.08 [-0.02 – 0.18]	.10	0.08 [-0.01 – 0.18]	.10
2. Composite Score			0.21 [0.12 – 0.31]	<.001		
3. Visuo-spatial processing					0.07 [-0.05 – 0.19]	.28
4. Quantitative knowledge					0.01 [-0.10 – 0.13]	.82
5. ATC Work Sample					0.08 [-0.02 – 0.19]	.12
6. Pilot Work Sample					0.13 [0.01 – 0.24]	.03
$R^2$	.01		.05		.06	
$\Delta R^2$			.04 ( $p < .001$ )		.05 ( $p < .001$ )	
AIC	1069		1053		1058	
RMSE	0.92		0.91		0.90	
Cross-validation RMSE	0.93		0.92		0.82	

### Unit Training Outcome

Table 4 shows the results of the logistic regression analyses conducted to predict the unit training outcome. In Model 1, explaining none of the variance, academic performance was not associated with unit training outcome. In Model 2a, composite score was added to the model and accounted for 6% of the variance in the unit training outcome above and beyond academic performance ( $p < .001$ ). The association between the composite score and the criterion was significant ( $OR = 2.89, p < .001$ ). Thus, evidence is found to support **H3**. In Model 2b, all cognitive abilities scores were introduced and accounted for 9% of incremental variance above and beyond academic performance ( $p = .001$ ). The pilot multitasking work sample was significantly associated with unit training outcome ( $OR = 1.98, p = .01$ ). The best model in terms of AIC was Model 2a ( $139 < 142$ ), the one with only the composite score. However, Model 2b has a better explained variance than Model 2a, probably due to the greater number of variables. For Model 2a and Model 2b, AUC and cross-validation AUC were equivalent, generalizability of both models can be expected contrary to Model 1.

**Table 4.**

*Logistic regression models for unit training outcome.*

Predictors	<i>Model 1</i>		<i>Model 2a</i>		<i>Model 2b</i>	
	OR	<i>p</i>	OR	<i>p</i>	OR	<i>p</i>
1. Academic performance	0.93 [0.60 – 1.46]	.75	0.91 [0.58 – 1.45]	.70	0.90 [0.56 – 1.45]	.65
2. Composite Score			2.89 [1.67 – 5.40]	<.001		
3. Visuo-spatial processing					1.04 [0.55 – 2.05]	.89
4. Quantitative knowledge					1.34 [0.77 – 2.37]	.30
5. ATC Work Sample					1.40 [0.85 – 2.31]	.18
6. Pilot Work Sample					1.98 [1.20 – 3.33]	.01
<i>R</i> <sup>2</sup>	.00		.06		.09	
$\Delta R^2$			.06 ( $p < .001$ )		.09 ( $p = .001$ )	
AIC	153		139		142	
AUC	.49		.75		.75	
Cross-validation AUC	.43		.73		.73	

### Optimal cut-off and operational impact

Based on the best model found for unit training outcome (Model 2a, the one with the composite score), an optimal cut-off value of -0.63 was computed. In other words, all participants with a composite score greater than -0.63 had a greater probability of success during unit training compared to those with a composite score lesser than -0.63. By applying this cutoff to our sample, the success rate increased from 92.4% to 95.3% (see Table 5), which would avoid three failures out of 100 ATCO trainees during unit training. Success rate among the potentially non-selected applicants would have been 82.2%, which is much lower than actual success rate. However, this highlights also the relatively high proportion of false negatives.

**Table 5**

*Contingency table after application of optimal cutoff.*

	FAILURE	SUCCESS
REJECT	11 (4%)	51 (18.4%)
SELECT	10 (3.6%)	205 (74%)

*Note.* Rows denote hypothetical selection outcome based on the threshold of -0.63 with Model 2a; columns denote actual unit training outcomes.

Supplementary analyses were carried out on a subset of en-route participants to understand the profile of false negatives; the majority of our participants being assigned to five en-route centers. Training duration was significantly higher ( $p < .01$ ) for the false negative (successful but rejected) participants compared to other participants in the same center. False negative participants required three months longer ( $M = 86$  days,  $SD = 211$ ,  $N = 35$ ) than the average compared to other participants ( $M = -1$  day,  $SD = 185$ ,  $N = 135$ ).

### Discussion

The objective of this longitudinal study was to test the incremental validity of cognitive abilities to predict ATCO training outcome for trainees pre-selected for their scientific academic proficiency. Indeed, 414 French ATCO students participated to a psychotechnical assessment with an experimental design, no

psychotechnical assessment being done during French ATCO selection. At the beginning of the training, they completed a psychotechnical assessment that included tests measuring visuospatial processing, quantitative knowledge, short-term memory, and processing speed. These specific cognitive abilities have been selected based on job analysis and meta-analysis. Then, initial training and unit training outcomes were collected. Psychotechnical predictors and academic performance were included in logistic regression to predict training outcomes. Our hypotheses were that cognitive predictors would have incremental validity compared to the current selection process (academic performance) for initial training and unit training outcomes. Our results provide evidence to support our hypotheses and demonstrate that, even for scientifically proficient applicants, cognitive abilities exhibited incremental validity beyond academic performance to predict both initial and unit training. Indeed, explained variance of training outcome was low (between 3% and 9%) but significantly increased with the addition of a cognitive abilities' composite score to the baseline regression model comprising only academic performance. This explained variance is close to the one provided by the meta-analysis of Sackett et al. (2024), which showed a low-to-medium correlation of general cognitive ability with job performance in the 21st century ( $r = .22$ ).

Cognitive abilities are known to predict ATCO training success (Martinussen et al., 2000; Mouratille et al., 2022) but students included in these meta-analyses were typically selected based on their performance in cognitive ability assessments and possession of at least a high school diploma. The present regression models' accuracies were equivalent to those of previously published validation studies. For example, 7% of variance were explained by cognitive ability assessment in a US sample of 260 candidates (Byrne & Broach, 2017) or 6% in a European sample of 225 candidates (FEAST, 2019). In the present study, as cognitive ability tests were not used to select ATCO trainees, data were not subject to range restriction concerning cognitive abilities among the population of interest. Therefore, our findings indicated that cognitive ability differences among scientific proficient applicants were predictive of a complex training such as air traffic control. This generalizes findings reported with less proficient applicants.

The composite score emerged as the best predictor of the practical grades during the initial and unit training (**H2** and **H3**). These findings were consistent with those highlighted by the meta-analyses (Martinussen et al., 2000; Mouratille et al., 2022) and generalize them to scientific proficient students. Indeed, cognitive composite scores may reflect executive functions (Agelink van Rentergem et al., 2020; Salthouse, 2005) and their incremental predictive validity highlights their importance for cognitively demanding jobs. In other words, among scientific proficient students, individual differences in executive functions exist and are significant predictors of success in highly demanding positions. In addition, when individual variables were introduced, only multitasking work sample tests were significant. A similar result was reported by Carretta and King (2008) with a low-fidelity ATC work sample test having incremental validity over the Armed Services Vocational Aptitude Battery. Our work sample tests required multitasking

abilities which are strongly influenced by executive functions (Himi et al, 2019). Moreover, individual differences in multitasking are related to working memory and attention control (Redick et al., 2016; Cowan et al., 2024) and already have been proven to predict job performance (Damos, 1993).

Based on Cattell-Horn-Carroll theory (Schneider & McGrew, 2018), short-term working memory (*Gwm*) is supposed to be the broad cognitive ability underlying multitasking (Himi et al., 2023; Vergauwe, von Bastian et al., 2022). Indeed, working memory is ‘the ability to maintain and manipulate information in active attention’ (Schneider & McGrew, 2018, p.). The main related narrow ability is attentional control, that is the flexible control of attention to focus on relevant stimuli while ignoring distractions (Schneider & McGrew, 2018). It enables goal pursuit despite distractions and maintains focus amid competing thoughts (Burgoyne & Engle, 2020). During practical training and unit training, students engage in realistic scenarios, either through simulations or real-life situations. A good performance during these realistic situations requires the handling of multiple tasks simultaneously. An ATCO can manage dozens of aircraft at the same time and must have an efficient working memory and attentional control. Attentional control is known to be critical in ATC (Redding et al., 1992) and, in a more general way, has been identified as a fundamental element underlying individual differences in cognitive abilities (Mashburn et al., 2024). Indeed, large effects were measured for working memory and work sample predictors in both meta-analyses, the same is observed in pilot selection (ALMamari & Traynor, 2019). In the present study, medium effects were measured for both multitasking work samples when analyzed individually.

The cognitive composite score was also significantly associated with the theoretical grades criterion (**H1**), but the model composed of all individual variables was more accurate. The best variable of this model was academic performance. In other words, a candidate with a high initial scientific knowledge is more likely to achieve high grades on theoretical assessments during the initial training. Moreover, a significant relationship between visuo-spatial processing and theoretical grades was observed. This ability has proven to be predictive of success in ‘Science, Technology, Engineering and Mathematics’ (STEM) learning (Buckley et al., 2018). Indeed, during the ATCO initial training, STEM courses are taught and need to be mastered to achieve high grades on theoretical assessments.

Quantitative knowledge had no predictive value in any model whereas this cognitive ability is a strong predictor in both ATCO meta-analyses and pilot selection (Johnson et al., 2017). This unexpected result can be explained by the specificity of our restricted sample. Indeed, scientific proficient students have already proven their competencies in quantitative knowledge. Thus, assessing this cognitive ability does not add incremental validity for this specific population of candidates.

Practical applications of present findings promote the use of cognitive ability testing, specifically through complex dynamical tests, even for applicants that have been pre-selected on their academic scientific proficiency. The failure rate could be reduced if pre-selected applicants were additionally evaluated based on their cognitive performance. This would reduce training costs, turnover rate and potentially safety concerns. However, in selection psychology even though false positives need to be lowest, false negatives must also be kept at an acceptable level. Indeed, especially when the pool of potential candidates is limited, there may be a preference for minimizing the risk of mistakenly rejecting candidates who could have succeeded. Our findings indicate a relatively high proportion of false negatives (82.2% of rejected applicants) while the proportion of false positive was quite low (4.6%). This failure rate would be particularly low compared to other ATCO schools (Langr et al., 2019). Therefore, if the pool of applicants is sufficiently large, selecting scientific proficient students with the highest cognitive ability performance assessed through complex multitask environments is a beneficial approach.

Several limitations are present in this study. The sample was relatively small, however, our sample size was equivalent to those used in studies such as Pecena et al. (2013) and Carretta and King (2008). Nevertheless, the data collection for the ATCO trainees who are still in unit training, would enable to confirm present findings even if cross-validation results provide evidence to support the hypothesis that the models can be generalized. Moreover, the use of an unrestricted sample regarding academic training performance, such as in Hoermann et al. (2018), would be the most suited research approach but is unfortunately not realistic, given the cost of ATCO training. In this case, data may be corrected for range restriction (Pfaffel et al, 2016). However, data from all candidates were not available due to GDPR concerns. In addition, since this study focused on the incremental validity of cognitive abilities in pre-selected candidates, the range-restriction correction was not necessary. Another limitation concerns the ATC work sample test and his reliability. The value of 0.67 is marginal and could be due to the innovative nature of the test. This is the first time they have “controlled” airplanes and the difficulty levels are perhaps too different. However, the other work sample test has a good reliability and has a better correlation. It could be interesting to make ATC work sample test more reliable, with a potential gain in predictive validity.

Another limitation is related to the training criteria. Thanks to European ATCO licensing, training criteria are standardized in all European countries. In the present study, ATCO trainees were successful or not based on this standardized training criteria. Training criteria were assessed by two assessors trained to increase inter-rater reliability, although it is recognized that this reliability may not be optimal, even in training settings (Speer et al., 2023; Zhou et al., 2024b). Nevertheless, as their operational performance

remains unknown, job performance could be a more relevant criterion (Allen et al., 2023). A last limitation could be the assessment context, which is different from an actual selection process. Indeed, applicants typically train themselves to increase their likelihood to be selected. However, this training is known to increase cognitive ability test scores (Hausknecht et al., 2002). This gain in score may have a positive (Hermes et al., 2023), or a negative impact (Matton et al., 2009) on the validity of the test scores.

## **Conclusion**

This study establishes that cognitive ability assessments (particularly working memory, processing speed, and executive control functions) significantly predict air traffic controller training success, even among scientifically proficient candidates. This is the first longitudinal study focusing on this specific population and including initial and unit training outcomes. Crucially, individual differences in these cognitive abilities explain significant variance in complex skill acquisition beyond domain-specific knowledge, demonstrating their valuable role in high-stakes multitasking environments. These findings confirm the incremental validity of cognitive assessment over knowledge-based selection across diverse populations. It is reasonable to assume that emerging developments in air traffic management, such as the integration of unmanned aircraft systems and multiple remote tower operations, will further heighten the importance of these cognitive functions.

## **Funding**

This work was supported by a Ph.D. fellowship from the French National School of Civil Aviation (ENAC), France.

## **Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- Ackerman, P. L., & Beier, M. E. (2018). 13 Methods for Studying the Structure of Expertise: Psychometric Approaches. *The Cambridge Handbook of Expertise and Expert Performance*, 213. <https://doi.org/10.1017/9781316480748.013>
- Agelink van Rentergem, J. A., de Vent, N. R., Schmand, B. A., Murre, J. M., Staaks, J. P., ANDI Consortium, & Huizenga, H. M. (2020). The factor structure of cognitive functioning in cognitively healthy participants: A meta-analysis and meta-analysis of individual participant data. *Neuropsychology Review*, 30(1), 51–96. <https://doi.org/10.1007/s11065-019-09423-6>
- Allen, M., Russell, T., Ford, L., Carretta, T., Lee, A., & Kirkendall, C. (2023). Identification and evaluation of criterion measurement methods. *Military Psychology*, 35(4), 308–320. <https://doi.org/10.1080/08995605.2022.2050165>
- ALMamari, K., & Traynor, A. (2019). Multiple test batteries as predictors for pilot performance: A meta-analytic investigation. *International Journal of Selection and Assessment*, 27(4), 337–356. <https://doi.org/10.1111/ijsa.12258>
- Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, 100(2), 431–449. <https://doi.org/10.1037/a0038047>
- Breit, M., Scherrer, V., Tucker-Drob, E. M., & Preckel, F. (2024). The stability of cognitive abilities: A meta-analytic review of longitudinal studies. *Psychological Bulletin*. 150(4), 399–439. <https://doi.org/10.1037/bul0000425>
- Broach, D. (2013). *Selection of the next generation of air traffic control specialists: Aptitude requirements for the air traffic control tower cab in 2018* (No. DOT/FAA/AM-13/5). Civil Aerospace Medical Institute.
- Broach, D., Byrne, C. L., Manning, C. A., Pierce, L., McCauley, D., & Bleckley, M. K. (2013). *The validity of the air traffic selection and training (AT-SAT) test battery in operational use* (No. DOT/FAA/AM-13/3). Civil Aerospace Medical Institute.
- Brown, M. I., Wai, J., & Chabris, C. F. (2021). Can you ever be too smart for your own good? Comparing linear and nonlinear effects of cognitive ability on life outcomes. *Perspectives on Psychological Science*, 16(6), 1337–1359. <https://doi.org/10.1177/1745691620964122>

- Buckley, J., Seery, N., & Canty, D. (2018). A heuristic framework of spatial ability: A review and synthesis of spatial factor literature to support its translation into STEM education. *Educational Psychology Review*, 30(3), 947–972.  
<https://doi.org/10.1007/s10648-018-9432-z>
- Burgoyne, A. P., & Engle, R. W. (2020). Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science*, 29(6), 624–630.  
<https://doi.org/10.1177/0963721420969371>
- Byrne, C. L., & Broach, D. (2017). The utility of aptitude in the placement of new air traffic controllers. In *Advances in Aviation Psychology, Volume 2* (pp. 87–103). Routledge.
- Carretta, T. R., & King, R. E. (2008). Improved military air traffic controller selection methods as measured by subsequent training performance. *Aviation, Space, and Environmental Medicine*, 79(1), 36–43. <https://doi.org/10.3357/ASEM.2166.2008>
- Clauser, B. E., & Bunch, M. B. (2021). *The History of Educational Measurement: Key Advancements in Theory, Policy, and Practice*. Routledge.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.
- Cowan, N., Bao, C., Bishop-Chrzanowski, B. M., Costa, A. N., Greene, N. R., Guitard, D., ... Ünal, Z. E. (2024). The relation between attention and memory. *Annual Review of Psychology*, 75, 183–214. <https://doi.org/10.1146/annurev-psych-040723-012736>
- Damos, D. L. (1993). Using meta-analysis to compare the predictive validity of single-and multiple-task measures to flight performance. *Human Factors*, 35(4), 615–628.  
<https://doi.org/10.1177/001872089303500403>
- Dehn, D. M., & Damitz, M. (2022). Do personality traits predict training performance? The use of a personality questionnaire for air traffic controller selection. *Aviation Psychology and Applied Human Factors*. <https://doi.org/10.1027/2192-0923/a000222>
- De Rooij, M., & Weeda, W. (2020). Cross-validation: A method every psychologist should know. *Advances in Methods and Practices in Psychological Science*, 3(2), 248-263.  
<https://doi.org/10.1177/2515245919898466>
- DuBois, P. H. (1947). *The classification program*. US Government Printing Office.
- Eaglestone, J., Damos, D., Hörmann, H., Stadler, K., & Wium, J. (2022). *Selection in Aviation : A European Association for Aviation Psychology Report*. Retrieved from <https://www.eaap.net/get/files/10022.pdf>
- Eurocontrol. (2022). *Network Operations Report 2021*. Retrieved from <https://www.eurocontrol.int/sites/default/files/2022-04/eurocontrol-annual-nor-2021-main-reportforconsultation.pdf>

- Federal Aviation Administration. (2014). *Hearing before the subcommittee on Financial and Contracting Oversight of the Senate Committee on Homeland Security and Government Affairs, 113th Congress, 2<sup>nd</sup> Session, January 14, 2014*. Retrieved from <https://www.govinfo.gov/content/pkg/CHRG-113shrg88271/html/CHRG-113shrg88271.htm>
- FAA. (2021). *FAA Aerospace Forecast 2021-2041*. Retrieved from [https://www.faa.gov/data%7B\\_%7Dresearch/aviation/aerospace%7B\\_%7Dforecasts/media/Forecast%7B\\_%7DHighlights.pdf](https://www.faa.gov/data%7B_%7Dresearch/aviation/aerospace%7B_%7Dforecasts/media/Forecast%7B_%7DHighlights.pdf)
- Confidential document FEAST. (2019). *Validation of the FEAST MULTI-PASS Test*. Brussels (Belgium): Eurocontrol.
- FEAST. (2024). Retrieved October 1, 2024, from <https://feast-info.eurocontrol.int>
- Goeters, K.-M., Maschke, P., & Eißfeldt, H. (2004). Ability requirements in core aviation professions: Job analysis of airline pilots and air traffic controllers. *Aviation Psychology: Practice and Research*, 99–119.
- Hausknecht, J. P., Trevor, C. O., & Farr, J. L. (2002). Retaking ability tests in a selection setting: Implications for practice effects, training performance, and turnover. *Journal of Applied Psychology*, 87(2), 243. <https://doi.org/10.1037/0021-9010.87.2.243>
- Hermes, M., Maier, J., Mittelstädt, J., Albers, F., Huelmann, G., & Stelling, D. (2023). Computer-based training and repeated test performance: Increasing assessment fairness instead of retest effects. *European Journal of Work and Organizational Psychology*, 32(3), 450–459. <https://doi.org/10.1080/1359432X.2023.2193692>
- Himi, S. A., Volberg, G., Bühner, M., & Hilbert, S. (2023). Individual differences in everyday multitasking behavior and its relation to cognition and personality. *Psychological Research*, 87(3), 655–685. <https://doi.org/10.1007/s00426-022-01700-z>
- Himi, S. A., Bühner, M., Schwaighofer, M., Klapetek, A., & Hilbert, S. (2019). Multitasking behavior and its related constructs: Executive functions, working memory capacity, relational integration, and divided attention. *Cognition*, 189, 275–298. <https://doi.org/10.1016/j.cognition.2019.04.010>
- Hoermann, H.-J., Noser, P., & Stelling, D. (2018). *A psychometrician's dream: Test validities in an unrestricted sample of future Swiss Air Force pilots*. Presented at the 33rd Conference of the European Association for Aviation Psychology (EAAP).
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96(1), 72. <https://doi.org/10.1037/0033-2909.96.1.72>
- Imbert, J.-P., Hodgetts, H. M., Parise, R., Vachon, F., & Tremblay, S. (2014). *The LABY Microworld: A Platform for Research, Training and System Engineering*. 58, 1038–1042. SAGE Publications Sage CA: Los Angeles, CA.

- Johnson, J. F., Barron, L. G., Carretta, T. R., & Rose, M. R. (2017). Predictive validity of spatial ability and perceptual speed tests for aviator training. *The International Journal of Aerospace Psychology*, 27(3-4), 109–120.  
<https://doi.org/10.1080/24721840.2018.1442222>
- Krolik, A. (2025). 285 of 313 air traffic control facilities are understaffed. *New York Times*. Retrieved 2025, from <https://www.nytimes.com/2025/01/31/business/air-traffic-controllers-understaffed.html>.
- Kulikowski, K., Cucina, J. M., Hayes, T. L., Oh, I.-S., Ones, D., & Viswesvaran, C. (2025). In defence of cognitive ability testing: Affirming the evidence for its continued use in personnel selection and admission decisions. *Journal of Occupational and Organizational Psychology*, 98, e70048.  
<https://doi.org/10.1111/joop.70048>
- Lancaster, B. P. (1999). Defining and interpreting suppressor effects: Advantages and limitations. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 139-148). JAI
- Lang, J. W., & Kell, H. J. (2020). General mental ability and specific abilities: Their relative importance for extrinsic career success. *Journal of Applied Psychology*, 105(9), 1047.  
<https://doi.org/10.1037/apl0000472>
- Langr, D., Ploch, J., Rudolf, P., Szabo, S., & Vajdova, I. (2019). Selection process as a key human aspect in air traffic control. *Transport Problems*, 14(3), 75–84.  
<https://doi.org/10.20858/tp.2019.14.3.7>
- Martinussen, M., Jenssen, M., & Joner, A. (2000). *Selection of air traffic controllers: Some preliminary findings from a meta-analysis of validation studies*. Presented at the Proceedings of the 24th (European Association for Aviation Psychology) Conference.
- Mashburn, C. A., Barnett, M. K., & Engle, R. W. (2024). Processing speed and executive attention as causes of intelligence. *Psychological Review*, 131(3), 664–694.  
<https://doi.org/10.1037/rev0000439>
- Matton, N., Paubel, P., Cegarra, J., & Raufaste, E. (2016). Differences in multitask resource reallocation after change in task values. *Human Factors*, 58(8), 1128–1142.  
<https://doi.org/10.1177/0018720816662543>
- Matton, N., Vautier, S., & Raufaste, E. (2009). Situational effects may account for gain scores in cognitive ability testing: A longitudinal SEM approach. *Intelligence*, 37(4), 412–421.  
<https://doi.org/10.1016/j.intell.2009.03.011>
- Ministry of Higher Education, Research and Innovation. (2013). *Note d'information Du SIES - Enseignement Supérieur, Recherche & Innovation*.
- Ministry of Higher Education, Research and Innovation. (2020). *L'accès à l'enseignement Supérieur*. Retrieved from <https://publication.enseignementsuprecherche>.

- [gouv.fr/eestr/FR/T943/l\\_acces\\_a\\_l\\_enseignement\\_superieur/](http://gouv.fr/eestr/FR/T943/l_acces_a_l_enseignement_superieur/)
- Mouratille, D., Amadiou, F., & Matton, N. (2022). A meta-analysis on air traffic controllers selection: Cognitive and non-cognitive predictors. *Journal of Vocational Behavior*, 138, 103769. <https://doi.org/10.1016/j.jvb.2022.103769>
- Nye, C. D., Ma, J., & Wee, S. (2022). Cognitive ability and job performance: Meta-analytic evidence for the validity of narrow cognitive abilities. *Journal of Business and Psychology*, 37(6), 1119–1139. <https://doi.org/10.1007/s10869-022-09796-1>
- Paton, L. W., McManus, I. C., Cheung, K. Y. F., Smith, D. T., & Tiffin, P. A. (2022). Can achievement at medical admission tests predict future performance in postgraduate clinical assessments? A UK-based national cohort study. *BMJ Open*, 12(2), e056129. <https://doi.org/10.1136/bmjopen-2021-056129>
- Pfaffel, A., Kollmayer, M., Schober, B., & Spiel, C. (2016) A missing data approach to correct for direct and indirect range restrictions with a dichotomous criterion: A simulation study. *PLoS one*, 11(3), e0152330. <https://doi.org/10.1371/journal.pone.0152330>
- Pecena, Y., Keye, D., Conzelmann, K., Grasshoff, D., Maschke, P., Heintz, A., & Eißfeldt, H. (2013). Predictive validity of a selection procedure for air traffic controller trainees. *Aviation Psychology and Applied Human Factors*. <https://doi.org/10.1027/2192-0923/a000039>
- Pecena, Y., Gayraud, K., & Eißfeldt, H. (2019). Aiming high maximizing predictive validity of selection processes for air traffic controller candidates. In *Pilot Selection* (pp. 355-368). CRC Press.
- Redding, R. E., Cannon, J. R., & Seamster, T. L. (1992). *Expertise in air traffic control (ATC): What is it, and how can we train for it?* Proceedings of the Human Factors Society Annual Meeting, 36(17), 1326-1330. <https://doi.org/10.1518/107118192786749397>
- Redick, T. S., Shipstead, Z., Meier, M. E., Montroy, J. J., Hicks, K. L., Unsworth, N., ... Engle, R. W. (2016). Cognitive predictors of a common multitasking ability: Contributions from working memory, attention control, and fluid intelligence. *Journal of Experimental Psychology: General*, 145(11), 1473. <https://doi.org/10.1037/xge0000219>
- Ree, M. J., & Carretta, T. R. (2022). Thirty years of research on general and specific abilities: Still not much more than g. *Intelligence*, 91, 101617. <https://doi.org/10.1016/j.intell.2021.101617>
- Rios, J. A., Liu, O. L., & Bridgeman, B. (2014). Identifying low-effort examinees on student learning outcomes assessment: A comparison of two approaches. *New Directions for Institutional Research*, 2014(161), 69–82. <https://doi.org/10.1002/ir.20068>
- Roe, R. A. (2017). The design of selection systems: Context, principles, issues. *The Blackwell Handbook of Personnel Selection*, 73–97. <https://doi.org/10.1002/9781405164221.ch4>

- Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118–137.  
<https://doi.org/10.1016/j.intell.2015.09.002>
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist*, *63*(4), 215.  
<https://doi.org/10.1037/0003-066X.63.4.215>
- Sackett, P. R., Demeke, S., Bazian, I. M., Griebie, A. M., Priest, R., & Kuncel, N. R. (2024). A contemporary look at the relationship between general cognitive ability and job performance. *Journal of Applied Psychology*, *109*(5), 687-713.  
<https://doi.org/10.1037/apl0001159>
- Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, *107*(11), 2040.  
<https://doi.org/10.1037/apl0000994>
- Salgado, J. F., & Moscoso, S. (2019). Meta-analysis of the validity of general mental ability for five performance criteria: Hunter and Hunter (1984) revisited. *Frontiers in Psychology*, *10*, 2227.  
<https://doi.org/10.3389/fpsyg.2019.02227>
- Salgado, J. F., Viswesvaran, C., & Ones, D. S. (2001). Predictors used for personnel selection: An overview of constructs, methods and techniques. *Handbook of Industrial, Work and Organizational Psychology*, *1*, 165–199. <https://doi.org/10.4135/9781848608320.n10>
- Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology*, *19*(4), 532. <https://doi.org/10.1037/0894-4105.19.4.532>
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262.  
<https://doi.org/10.1037/0033-2909.124.2.262>
- Schneider, W. J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 73–163).
- Speer, A. B., Delacruz, A. Y., Wegmeyer, L. J., & Perrotta, J. (2023). Meta-analytical estimates of interrater reliability for direct supervisor performance ratings: Optimism under optimal measurement designs. *Journal of Applied Psychology*.  
<https://doi.org/10.1037/apl0001146>

- Vergauwe, E., von Bastian, C. C., Kostova, R., & Morey, C. C. (2022). Storage and processing in working memory: A single, domain-general resource explains multitasking. *Journal of Experimental Psychology: General*, *151*(2), 285.  
<https://doi.org/10.1037/xge0000895>
- Wang, X. (2014). Firth logistic regression for rare variant association tests. *Frontiers in Genetics*, *5*, 99197.  
<https://doi.org/10.3389/fgene.2014.00187>
- Watrin, L., Geiger, M., Levacher, J., Spinath, B., & Wilhelm, O. (2022). Development and initial validation of an admission test for Bachelor psychology studies. *Frontiers in Education*, *7*, 909818.  
<https://doi.org/10.3389/feduc.2022.909818>
- Wilson, M. A. (2007). A history of job analysis. *Historical Perspectives in Industrial and Organizational Psychology*, 219–241.
- Wium, J., & Eaglestone, J. (2022). Job Analysis for Air Traffic Control. *Aviation Psychology and Applied Human Factors*. <https://doi.org/10.1027/2192-0923/a000218>
- Wolfram, T. (2023). (Not just) Intelligence stratifies the occupational hierarchy: Ranking 360 professions by IQ and non-cognitive traits. *Intelligence*, *98*, 101755.  
<https://doi.org/10.1016/j.intell.2023.101755>
- Woo, S. E., LeBreton, J. M., Keith, M. G., & Tay, L. (2023). Bias, fairness, and validity in graduate-school admissions: A psychometric perspective. *Perspectives on Psychological Science*, *18*(1), 3–31.  
<https://doi.org/10.1177/17456916211055374>
- Woods, S. A., & Patterson, F. (2024). A critical review of the use of cognitive ability testing for selection into graduate and higher professional occupations. *Journal of Occupational and Organizational Psychology*, *97*(1), 253-272.  
<https://doi.org/10.1111/joop.12470>
- Zhou, Y., Kuncel, N. R., & Sackett, P. R. (2024a). Examining the Existence of Cognitive Thresholds in Highly Quantitative College Courses. *Journal of Intelligence*, *12*(4), 37.  
<https://doi.org/10.3390/jintelligence12040037>
- Zhou, Y., Sackett, P. R., Shen, W., & Beatty, A. S. (2024b). An updated meta-analysis of the interrater reliability of supervisory performance ratings. *Journal of Applied Psychology*.  
<https://doi.org/10.1037/apl0001174>
- Zwick, R. (2019). Fairness in measurement and selection: Statistical, philosophical, and public perspectives. *Educational Measurement: Issues and Practice*, *38*(4), 34–41.  
<https://doi.org/10.1111/emip.12299>