Selection and Pre-Employment Assessment in Aviation Security X-Ray Screening

Anton Bolfing
School of Applied Psychology
Universiy of Applied Sciences Northwestern Switzerland
CH-4600 Olten, Switzerland
and
Center for Adaptive Security Research
and Applications (CASRA)
Thurgauerstr. 39
8050 Zurich, Switzerland

Adrian Schwaninger
School of Applied Psychology
Universiy of Applied Sciences Northwestern Switzerland
CH-4600 Olten, Switzerland
and
Center for Adaptive Security Research
and Applications (CASRA)
Thurgauerstr. 39
8050 Zurich, Switzerland

Abstract—This document concentrates on how aviation security can be largely improved by selecting the right people for the x-ray screening task. X-ray images dispose of a different colour range; single objects superimpose each other independently of whether they are in front or behind of each other, and the single objects look very different from what they look like under natural light. Therefore, individual visual abilities such as mental rotation, figure-ground segregation, or visual search are very important for the correct interpretation of x-ray images. In this paper we briefly discuss job and task analysis, as well as cognitive task analysis in x-ray screening. We then discuss tests used in pre-employment assessment, including a more detailed presentation of the X-Ray Object Recognition Test (X-Ray ORT). Finally, a new project is presented in which a cognitive test battery is developed to measure important visual abilities needed in x-ray screening and related tasks.

Keywords - aviation security human factors, pre-employment assessment, visual abilities, selection tests

I. INTRODUCTION

Together with the tremendous growth of civil aviation in the last decades, the importance of aviation security has dramatically increased. Despite of large technological progress in x-ray screening equipment, the final decision whether a luggage is allowed to enter an aircraft or not is still made by a human operator. Human factors remain the essential key element in aviation security, which is consistent with covert test results that have sometimes shown substantial deficiencies at airport security controls.

When comparing x-ray images to everyday visual information, several differences are apparent. X-ray images dispose of a different colour range; single objects superimpose each other independently of whether they are in front or behind of each other and single objects look very different compared to what they look like under natural light. Therefore, individual visual abilities such as mental rotation, figure-ground segregation, or visual search are essential for a reliable and fast interpretation of x-ray images (Hardmeier & Schwaninger, 2008). In addition, several other factors need to be considered for an effective and reliable pre-employment assessment.

When testing people in pre-employment assessments it is important not to select them based on tests that are highly dependent on training but on tests that measure relatively stable abilities and aptitudes. Factors that should be assessed in pre-employment assessments should be those factors important for a certain job/task where large individual differences exist and which cannot easily be trained. Therefore, it needs to be investigated which tasks are important for the aviation security job and what relatively stable abilities are needed to successfully fulfill the job requirements. Although various tasks at a security checkpoint are briefly discussed in this article, we focus on the x-ray screening task because it still is the most important task in terms of abilities, aptitudes and training needed to achieve good performance.

In the following section, job and task analysis are briefly discussed. We then present tests that are currently being used in pre-employment assessment. Finally, current developments regarding a visual cognition test battery for pre-employment assessment are presented and discussed.

II. JOB AND TASK ANALYSES

Job and task analytic techniques are useful tools to understand the context in which the human operators work is taking place. A job and task analysis should be performed primarily in order to identify the important tasks and responsibilities of a specific job. The job and task overview then provides a clear description with which to define the most important tasks. These should further be analysed by means of a traditional or cognitive task analysis (CTA), depending on the task demands (Kirwan & Ainsworth, 1992; Crandall, Klein & Hoffmann, 2006).

Aviation security controls are conducted to ensure that no threat items or dangerous goods can enter the security restricted area or an aircraft. For the job analysis in cabin baggage screening (CBS), a subject-based and an observation-based data collection method was chosen. That is, inputs from
persons who were familiar with the task were collected, and observation methods were applied by professionals. A typical setting of an aviation security checkpoint is shown in figure 1. Each screener works at an assigned position for a defined period of time. Normally after 20 to 30 minutes, positions are changed. Each position is linked to a specific task. X-ray screening of passenger bags (see figure 1, position 2) is one of the most important tasks at the checkpoint, since missing a threat item in a passenger bag could have terrible consequences. Note that if a screener is not effective in the visual inspection of x-ray images and therefore sends too many harmless bags to be hand-searched, long waiting lines at the checkpoint will occur. Many objects look quite different in x-ray images than in reality. Thus, for the screening process, visual abilities and knowledge of the visual appearance of threat items in x-ray images are important determinants in achieving a certain level of security without sacrificing efficiency.

If the screener decides that a piece of baggage has to be hand searched, another screener (see figure 1, position 3) is responsible for hand searching the bag (although at some airports screeners change position when an x-ray image is sent to bag search so that the same screener can hand search the bag). Passengers could also try to bring prohibited items into the security restricted area and the airplane by wearing them on their body. Therefore, at minimum one female and one male screener are responsible for the body check (figure 1, position 4). At most airports, the body check is done by means of a metal detector and sometimes an additional manual body search. In coming years, this manual body search might be replaced by new technology (e.g., millimetre wave and backscatter x-ray systems). Millimetre wave technology allows the scanning of people for the presence of threat objects. As clothing and other organic materials are translucent, an image of the passenger can be provided, which can then be interpreted.

Another position is in front of the X-ray machine (figure 1, position 1). The screener informs passengers about the security check and places their bags on the conveyor belt to ensure a smooth and effective screening procedure (it ensures also that the distance between two bags is large enough and thus enough time is provided to the operator who has to interpret the x-ray image).

For nearly all positions in the CBS area, dealing with passengers is important. As a result, communication between aviation security screeners and passengers can be assumed to play a key role in ensuring an efficient workflow. Hence, language and communication skills as well as customer service skills were defined as basic job requirements. Moreover, a crew always consists of several screeners who have to work as an efficient team, not only during normal operations but especially in stressful situations. This factor becomes even more important taking into consideration the fact that screeners often are randomly assigned to a crew and especially at bigger airports do not know each other well. Furthermore, passengers are often not pleased to have to pass through security control, and therefore screeners have to be very patient and need the ability to cope with negative feedback even if they do their job very well.

In summary, the job analysis revealed x-ray screening, baggage search, body search, dealing with passengers, teamwork, and coping with negative feedback to be important for airport security screeners working in the CBS area.

III. COGNITIVE TASK ANALYSIS IN X-RAY SCREENING

Based on the job analysis, the abilities, aptitudes, and knowledge that are needed to perform the tasks proficiently should be identified. Depending on the task, a traditional
(behavioural) task analysis or a cognitive task analysis (CTA) can be applied. Whereas the traditional task analysis focuses mainly on non-critical procedural tasks, the CTA identifies and describes cognitive elements, processes such as decision making, problem solving, and so on, as well as the knowledge and skills that are required for similar job components. Behavioural task analysis describes the task in terms of time spent, criticality, and frequency. In contrast, CTA should be used for high-performance tasks that require large amounts of knowledge or information, significant decision making or problem solving, heavy workload or time pressure, multitasking, situations that change substantially, or considerable amounts of teamwork.

Generally, these cognitive processes are more difficult to study because they are not directly observable. Traditional task analysis and CTA can also be used as complementary tools. Usually traditional task analyses are used to specify the basic job tasks and precede the CTA. Quite often a CTA includes a comparison of novices and experts, to find out which skills and knowledge are domain-specific. These findings can help to identify selection criteria for required skills and define the training sessions needed to acquire job specific knowledge and procedures.

IV. TESTS USED IN PRE-EMPLOYMENT ASSESSMENT

The Ishihara Colour Vision Test, the X-Ray Object Recognition Test, language and communication tests as well as medical examinations are currently in use in several European countries. In the following subsections, we focus on tests related to visual information processing used in x-ray image interpretation.

1) Ishihara Colour Vision Test: Since state-of-the-art x-ray imaging equipment disposes of sophisticated colour coding of the x-rayed materials highly facilitating x-ray image interpretation colour vision as an absolutely mandatory precondition of aviation security officers. The most established colour vision test is the so-called Ishihara Colour Test first published in 1917 (Ishihara, 1917; see figure 2). The test consists of a series of so-called Ishihara plates. Common plates include a circle of dots in shades of green and light blues with a figure differentiated in shades of brown, or a circle of dots in shades of red, orange and yellow with a figure in shades of green; both testing for red-green and blue-yellow colour blindness.

It is highly recommended to conduct this test using the original Ishihara plates, since computer versions are very vulnerable to display colours inaccurately.

2) Visual acuity test: Normal or corrected to normal vision regarding visual acuity is another prerequisite for becoming a screener. Reliably seeing fine detail is essential for x-ray image interpretation (e.g. to detect thin wires of an improvised explosive device). Thus it is highly recommended to use a visual acuity test as part of pre-employment assessment.

3) X-Ray Object Recognition Test: To specifically assess visual abilities required in x-ray image interpretation, the X-Ray Object Recognition Test (X-Ray ORT) was developed.

This test measures reliably visual abilities that are relatively independent of visual knowledge acquired through training and experience. It only contains guns and knives as prohibited objects, i.e. shapes that are expected to be known from everyday life, or at least from every-day multimedia entertainment. Knowledge of what more specific prohibited items like IEDs (improvised explosive devices) or tasers (electric shock devices) look like in x-ray images is not required to successfully solve this test. Because the meaning of colour information in x-ray images is not available to novices, x-ray images in the X-Ray ORT are in greyscale.

In the X-Ray ORT three specific image-based factors closely related to visual abilities are systematically manipulated: view difficulty, superposition and bag complexity. These image-based factors are known to highly influence threat detection performance. So do the related visual abilities from a human factors perspective. View difficulty is closely related to the visual ability to mentally rotate objects. Superposition is closely related to the ability to differentiate between overlaying shapes of different objects in an x-ray image.
(figure-ground segregation) and bag complexity is related to the ability to visually search for specific patterns in complex images.

The X-Ray ORT consists of 256 x-ray images of passenger bags that must be judged for whether they are OK or NOT OK to enter an aircraft. One half of all images contain a threat item, those of the other half do not. The 128 threat trials of the test result from the following combinations: 8 guns and 8 knives are being projected into bags with low and high bag complexity (BC) with low and high superposition (SP) and in an easy and in a difficult view (view difficulty, VD) each. Thus, all 16 threat exemplars are shown in all 8 factor combinations (2 VD x 2 SP x 2 BC; see figure 4). At a busy airport security checkpoint, screeners often have only a few seconds to decide whether an x-ray image is OK or not. In order to take this into account, the maximum display duration of x-ray images in the X-Ray ORT is limited to 4 seconds. The X-Ray ORT is a reliable, valid and standardized test which has been documented in several scientific publications (Hardmeier, Hofer, & Schwaninger, 2006a; Hardmeier, Hofer and Schwaninger, 2005).

The X-Ray ORT is a very reliable test, Cronbachs alpha values range from .887 to .966 for professional screeners and from .907 to .970 for novices. Split-half reliability values of the X-Ray ORT range from .781 to .904 for professionals and from .778 to .939 for novices. Large effects of bag complexity, superposition and viewpoint could be shown for aviation security screeners and novices and support high internal validity. Furthermore, convergent and discriminant validity could be shown based on 453 screeners who were selected with the old selection procedure, correlating results in the X-Ray ORT with results in a prohibited items test ($r = .61$, $p < .001$) and results in a computer-based questionnaire ($r = .27$, $p < .001$), respectively.

To test whether the X-Ray ORT is a useful pre-employment assessment tool, detection performance of screeners selected without the X-Ray ORT and screeners who were selected after having passed the X-Ray ORT was compared one year later in an x-ray image interpretation certification test (Hardmeier, Hofer, & Schwaninger, 2006a). All screeners who were selected with the X-Ray ORT had completed a classroom training and about one year of working experience before taking the certification test. The detection performance was calculated using the "nonparametric" detection performance measure A (see Pollack & Norman, 1964; MacMillan & Creelman, 1991; Hofer & Schwaninger, 2004). A takes into account the hit rate (i.e. bags containing a prohibited item judged as NOT OK) as well as the false alarm rate (i.e. harmless bags judged as NOT OK).

In order to investigate whether the X-Ray ORT is a valuable tool for pre-employment assessment, the mean detection performance of both groups in the certification test was compared (see figure 5). A significant difference in detection performance between screeners selected without the X-Ray ORT and screeners selected with the X-Ray ORT was shown.

The job applicants who were selected with the X-Ray ORT were significantly better in detecting prohibited items in x-ray images, $t(552) = 14.51$, $p < .001$.

4) **Cognitive Test Battery:** Although the X-Ray ORT itself assesses important cognitive visual abilities required in aviation security x-ray image interpretation very reliably and with a high external validity, the X-Ray ORT is still not completely independent from experience in x-ray image interpretation, since its tasks are all about judging x-ray images. Further, the important cognitive visual abilities are not assessed separately in the X-Ray ORT. In order to get a deeper understanding of the screener’s visual cognitive abilities we are currently developing a test battery containing a series of very specific tests assessing different cognitive functions/abilities separately. To date, the implemented cognitive functions are strained to visual abilities. Please refer to section V for short description of the subtests currently implemented in the visual cognition test battery. The introduced subtests are still under development, they are currently being optimised for reliability, validity and predictive power regarding x-ray image interpretation competency later on the job. This introduced visual cognition test battery shall become the core element of a larger cognitive test battery exceeding the mere visual aspect of the aviation security offer’s job. We intend to expand the tests for other cognitive aspects such as for example concentrativeness, vigilance, alertness and other aspects that are not yet contained explicitly.

5) **Pre-employment assessment tests aside from the vision context:** The field of activity of aviation security officers is not limited to x-ray image interpretation. Although this document concentrates on the visual cognition aspect, it is important to note that other human characteristics are worth being assessed in a pre-employment assessment procedure. X-ray image inspection can be a fatiguing job. Fortunately, aviation security screeners have additional tasks at most security checkpoints. Examples are instructing passengers and placing bags on the x-ray belt, conducting manual search of luggage or conduct an additional test using a hand held metal detector. All these tasks need specific skills. Aviation security officers must have substantial social and communication skills especially
when dealing with certain passengers. Furthermore, they must have a certain constitution since they work about and must lift luggage and search it. Finally, the employer is interested in employing people with high degrees of conscientiousness, integrity, discretion, secrecy, work ethic, and service orientation. All these aspects should be taken in account when selecting new personnel for becoming an aviation security screener.

V. Visual Cognition Test Battery

As explained in section IV-4 we are currently developing a visual cognition test battery. The goal of this test battery is to reliably and efficiently assess visual abilities which are important for x-ray image interpretation. Two aspects, efficiency and modularity are very important in pre-employment assessment. Efficiency in terms of how long does a test take, is important in terms of costs and can be achieved by construction of short but reliable tests. The great advantage of modularity is that test results can directly be interpreted as personal profiles providing information on the personal strengths and weaknesses.

The first visual cognitive functions to begin developing tests for are the three abilities connected to the three image based factors introduced in section IV-3, metal rotation (view difficulty), figure ground segregation (superposition) and visual search on complex backgrounds (bag complexity). Additionally, Hardmeier & Schwaninger, 2008 could show that 'visual' reasoning is another important factor in predicting x-ray image interpretation. In this paper they explored the predictive power of twelve well established psychological tests, mainly visual, concentration and reasoning tests, based on two populations of 169 participants in 2006 and 97 in 2007. In our current visual cognition test battery we tried to develop tests measuring the same underlying cognitive functions as these five established tests do.

In the following, the subtests implemented to date in our visual cognitive test battery are presented in short in the order they appear in our test battery. Unfortunately, there are no data available yet regarding the predictive power for x-ray image interpretation competency later on the job. First reliability analyses have been conducted and the tests are now being optimized according to item analyses including replacement of bad items by newly created items. Further, for each test we will evaluate how they are best analysed in terms of response time restriction and detection performance measures in regarding high predictive power for x-ray image interpretation competency later on the job.

A. Reasoning

The first subtest consists of 30 trials presenting a 3x3 matrix with eight cells containing abstract visual stimuli logically connected together. The bottom right cell is a placeholder (shaded grey in figures 6 and 7) for which the candidate’s task is to find the logically fitting element from the eight selection options (arranged in 2x4 matrix) at the bottom. In our current visual cognition test battery we tried to develop tests measuring the same underlying cognitive functions as these five established tests do.

In the following, the subtests implemented to date in our visual cognitive test battery are presented in short in the order they appear in our test battery. Unfortunately, there are no data available yet regarding the predictive power for x-ray image interpretation competency later on the job. First reliability analyses have been conducted and the tests are now being optimized according to item analyses including replacement of bad items by newly created items. Further, for each test we will evaluate how they are best analysed in terms of response time restriction and detection performance measures in regarding high predictive power for x-ray image interpretation competency later on the job.

B. Mental rotation

1) Planar rotation: The subtest consists of 32 trials displaying three identical x-ray images of threat items in different planar rotations. One of these threat items is additionally mirrored. The assessee’s tasks is to decide as fast as possible which one of the three items is mirrored. The optimal mode
of response time analysis is currently being investigated. Figures 8 and 9 show two examples of test trials.

2) 3D mental rotation: The third and fourth test of the current visual cognition test battery measures mental rotation regarding different axes. For each of the following three figures 10, 11 and 12, the candidate’s task consists in deciding whether the two figures are rotated (ROTIERT), hence identical or whether they are mirrored (GESPIEGELT), hence different. The third subtest of this test battery displaying two blue tubes consists of 30 trials. Figure 10 shows a trial with different tubes with the answer options.

Figure 11 shows two identical tubes that are just differently rotated.

Figure 12 shows a test trial of the fourth subtest, which consists in a slight variation of the third test displaying blue tubes. The fourth tests consists of 30 trials showing two different star-like geometrical objects. The candidate’s task is exactly the same as in the third test, namely to decide whether the two objects are identical (rotated) or not (mirrored).

C. General spatial imagination I

The fifth test of the test battery simply consists in correctly counting the number of planes of 26 three-dimensional upper case letters and to indicate it on a response bar at the bottom of the screen. This task will be analysed by response times. Again, the optimal mode of response time analysis is not fixed yet. Figure 13 shows an example of the task.

D. Visual search on a complex background

The sixth test in battery measures the ability to visual search for a known object on a very complex background. All 28 trials consist of chaotic patterns of lines crossing each other in 60° angles. All lines are coloured between a certain colour (red, green, blue, turquoise, violet, yellow or white) and black. The candidate’s task simply consists in finding a small black
E. General spatial imagination II

The seventh and last subtest of the battery also measures general spatial imagination. The candidate’s task is to mentally unfold a 3D object and to allocate it to one of three selection options as shown in the two example figures 16 and 17.

VI. CONCLUSION

Since not all job requirements for the aviation security officers are easily be acquired by instructions and since insufficient abilities cannot easily compensated by training, it is very important to select people for this job that already dispose of the most important required abilities. Therefore, a reliable, valid, standardized and sophisticated pre-employment assessment for aviation security officers is needed. The most expensive technology is of little value if the humans who operate them are not selected and trained appropriately. In this article different aspects were discussed which need to be taken into account for the development of pre-employment assessment procedures. Finally, a new visual cognition test battery especially designed to assess the most important visual abilities was presented that is currently under development. We expect this visual cognition test battery to become a very reliable and efficient tool for selecting the right people for the visually most demanding jobs in aviation security and other job areas.
Figure 17. Another example trial of the seventh test in the battery.

ACKNOWLEDGMENTS

The author would like to thank Sandra Iglesias, Silke Lengler and Alexei Fischer for their valuable work regarding the conceptualisation and creation of stimuli for the new visual cognition test battery.

REFERENCES