

The role of recurrent CBT for increasing aviation security screeners' visual knowledge and abilities needed in x-ray screening

Diana Hardmeier

University of Zurich, Department of Psychology, Visual Cognition Research Group (VICOREG),
Switzerland

Franziska Hofer and Adrian Schwaninger

University of Zurich, Department of Psychology, Visual Cognition Research Group (VICOREG),
Switzerland

Abstract

X-ray screeners have to know which items are prohibited and what they look like in x-ray images of passenger bags (knowledge-based factors). In addition, effective x-ray screening requires the abilities to cope with bag complexity, superposition by other objects, and rotation of objects (image-based factors). Knowledge-based factors are expected to be highly dependent on training whereas image-based factors are related to visual-cognitive abilities and aptitudes (Schwaninger, Hardmeier, & Hofer, 2005).

To test to what extend these two factors are influenced by training, 334 screeners took two x-ray screening tests before and after two years of recurrent computer-based training (CBT). The Prohibited Items Test (PIT) measures rather knowledge-based factors, the X-Ray Object Recognition Test (X-Ray ORT) image-based factors. The results showed indeed a much better detection performance in the PIT after two years of training. Thus, CBT can increase the knowledge of prohibited items and what they look like in x-ray images of passenger bags substantially. The increase in detection performance in the X-Ray ORT was much smaller. This indicates that image-based factors are indeed related to visual-cognitive abilities and aptitudes that can be increased by CBT less effectively. The implications for selection and training of x-ray screeners are discussed.

Introduction

Airport security has become very important in recent years. Since airports are confronted with new threat dimensions and a constantly increasing passenger flow, reliable and efficient detection of different threat items in x-ray images is an essential airport security task. During high passenger flow, screeners have only a few seconds to decide whether a bag is OK or whether it has to be hand-searched. Schwaninger, Hardmeier, and Hofer (2005) could show that detecting threat items in passenger bags includes both, knowledge-based and image-based factors. That is, a screener has to know which items are prohibited and what they look like in x-ray images of passenger bags (knowledge-based factors). There are many threat items that look quite different in reality than in an x-ray image, which makes them difficult to recognize without training (Schwaninger, 2005b). Another difficulty results from the fact that some threat items look quite similar to harmless objects. Again other threat items like improvised explosive devices (IEDs) are normally not seen at checkpoints and without specific training they are therefore rather difficult to recognize. These knowledge-based factors are expected to be highly dependent on training since results from object recognition studies show that one can only recognize shapes if they are similar to the ones encountered before and stored in visual memory (for an overview see Schwaninger, 2005a). However, the interpretation of x-ray images is much more complex and is also dependent on image-based factors such as bag complexity, superposition and rotation of the target object itself. If a bag is close-packed it is more difficult to detect a threat item within a short time because other objects can distract attention. Furthermore, the threat item can be superimposed by other objects in the bag, which can hamper detection performance as well. Moreover, if a threat item is shown in a rotated view it becomes harder to recognize it. A previous study from Schwaninger et al. (2005) showed large individual differences for novices and screeners regarding how well they can cope with such image-based factors.

The 4th International
Aviation Security Technology Symposium
November 27 - December 1, 2006

Furthermore, it could be assumed that these image-based factors are less dependent on training and more related to visual abilities and aptitudes which are relatively stable over time.

To measure knowledge-based and image-based factors in x-ray screening relatively independent of each other, two x-ray screening tests, the Prohibited Items Test (PIT) and the X-Ray Object Recognition Test (X-Ray ORT) were developed. The PIT measures rather knowledge-based factors in x-ray screening and therefore includes all kinds of prohibited items according to international prohibited items lists. On the other hand the X-Ray ORT includes only guns and knives in x-ray images, but shown in different viewpoints with low and high superposition and in bags with different complexity levels and therefore measures mainly the image-based factors viewpoint, superposition and bag complexity.

To test to what extent knowledge-based and image-based factors are influenced by training, 334 aviation security screeners took both x-ray screening tests before and after two years of recurrent computer-based training with X-Ray Tutor, which is an individually adaptive training system for x-ray screeners (see Schwaninger, 2004; Schwaninger & Hofer, 2004 for details).

Method

Participants

A total of 334 aviation security screeners (101 male and 233 female) between 23 and 62 years ($M = 46.71$, $SD = 8.37$) participated in this study. When taking the PIT and the X-Ray ORT the first time, all of them had a working experience between one and 25 years ($M = 7.54$, $SD = 5.13$). Between the first and the second measurement all screeners received two years of recurrent Computer Based Training (CBT) with X-Ray Tutor. Most screeners trained at least twice a week for 20 minutes.

Materials and Procedure

Prohibited Items Test (PIT)

The Prohibited Items Test (PIT) measures rather knowledge-based factors in x-ray image interpretation tasks and includes x-ray images of all kinds of prohibited items. All of them can be classified into the seven categories guns, sharp objects, hunt and blunt instruments, chemicals, highly inflammable substances, explosives and others according to ECAC, ICAO and EU prohibited items lists. The test contains a total of 160 x-ray images, half of them include one or more prohibited items, whereas the other 80 images are bags containing only harmless objects. 68 of the threat images contain exactly one prohibited item; the remaining 12 bags include two or three prohibited items¹. As this test was developed to measure rather knowledge-based factors, all image-based factors are kept relatively constant. That is, all prohibited items were shown in a bag with medium bag complexity, medium superposition and in an easy view.

The PIT is a computer based test and includes a self-explanatory instruction with six exercise trials to familiarize the participants with the test taking procedure. All x-ray images are shown for a maximum of 10 seconds on the screen. Then, participants have to decide whether the bag is OK (includes no prohibited item) or NOT OK (includes one or more prohibited items) by clicking on the respective button. Furthermore, they have to indicate how sure they are in their decision and to which of the seven categories the prohibited item belongs to². The test is divided into four blocks of trials. After each block screeners had the possibility to take a short break. Trials are counterbalanced across all four blocks and the order within a block is random. The test takes about 45 minutes (short breaks included).

A previous study with 453 x-ray screeners could show that the PIT is a reliable and valid instrument to measure knowledge-based factors in x-ray screening (Hardmeier, Hofer, and Schwaninger, 2006). Reliability was measured using Cronbach Alpha and Guttman split half reliabilities; the former ranging between .87 and .93 and the latter ranging between .87 and .94. Furthermore, convergent, discriminant and criterion-related validity support high validity of the PIT (for more details see Hardmeier et al., 2006).

¹ This was done to increase face validity of the test. Please note that only bags including one prohibited item were used for that data analysis.

² Please note that data analysis is based on the OK/NOT OK answers and not on the answer to which category a prohibited item belongs to.

X-Ray Object Recognition Test (X-Ray ORT)

The X-Ray Object Recognition Test (X-Ray ORT) measures rather image-based factors in x-ray screening and therefore includes only typical gun and knife shapes as threat objects. These are shown in bags with different complexity levels, more or less superimposed by other objects in the bag. Furthermore, each gun and each knife is shown in an easy and rotated view. Thus, the *X-Ray ORT* consists of a total of 256 trials: 2 threat categories (guns and knives) * 8 (exemplars) * 2 (bag complexities) * 2 (superpositions) * 2 (views) * 2 (harmless images vs. threat images). All images are shown in black and white so that this test can also be used for pre-employment assessment purposes where the meaning of color information as indicator for different materials is not known to novices.

The procedure is similar to the PIT. After a self-explanatory introduction, participants receive eight exercise trials with feedback. The test is also subdivided into four blocks. Trials are counterbalanced across all blocks and random within each block. Contrary to the PIT, images in the ORT are displayed for four seconds only and then participants answer whether the bag was OK or NOT OK. Again, at the end of each answer participants have to indicate how sure they are in their decision.

Detailed reliability and validity measures of this test can be found in Hardmeier, Hofer, and Schwaninger (2005). Overall, reliabilities with $>.89$ for Cronbach Alpha and $>.78$ for Guttman split half in the screener group are rather high. As well convergent, discriminant and criterion-related validity are given; for details see Hardmeier et al. (2005).

X-Ray Tutor – Individually adaptive computer based training system

X-Ray Tutor (XRT) is an individually adaptive training system to improve detection performance in x-ray screening. This CBT system creates individual training sessions adapted to each screener based on his learning history and thereby provides very effective and efficient training (for details see Schwaninger 2004). XRT CBS 2.0 Professional Edition includes a large image library with 25'140 fictional threat item (FTI) images depicting more than 500 different threat objects in many different viewpoints. XRT combines each FTI with an x-ray image of a passenger bag during the training session in real-time. The training software starts with easy views at the beginning of the training. Depending on the screeners' learning history, image difficulty is increased by choosing more difficult viewpoints, increasing bag complexity and superposition adapted to each screener and FTI.

During a training session the x-ray image is displayed for 15 seconds on the screen. Then, screeners have to press an OK or NOT OK button to indicate whether the bag is harmless or whether it has to be hand-searched. Immediate feedback is provided, i.e. whether a screener has correctly identified (hit) or missed the threat item (miss), whether he/she correctly rejected a harmless bag (correct rejection) or wrongly judged a harmless bag as being dangerous (false alarm). Furthermore, an information window showing the x-ray image of the threat item and a real photograph of it provide immediate detailed information about the threat item and its components in order to enhance perceptual learning. The effectiveness of X-Ray Tutor has been proven in several scientific studies, showing substantial increases of detection, less false alarms and faster response times (Schwaninger & Hofer, 2004; Ghylin, Drury, & Schwaninger, 2006).

Results

Test results are calculated using the detection performance measure d' , which takes the hit rate and the false alarm rate into account. The hit rate shows how often a bag containing a threat item was judged as being not ok, whereas the false alarm rate shows how often a harmless bag was wrongly judged as not ok (Green & Sweets, 1966).

Figure 1a shows the difference in detection performance after two years of recurrent CBT for both tests, the PIT and X-Ray ORT. An analysis of variance (ANOVA) with the within-participant factors test type (PIT, ORT) and measurement (first, second) was calculated using d' scores. There was a significant main effect of test type (PIT vs. ORT) $\eta^2 = .88$, $F(1, 333) = 2483.84$, $p < .001$, a significant main effect of measurement (first vs. second) $\eta^2 = .77$, $F(1, 333) = 1129.58$, $p < .001$ and a significant interaction of test type and measurement $\eta^2 = .34$, $F(1, 333) = 170.44$, $p < .001$. As can be seen, CBT had a much higher influence on detection performance in the PIT than in the X-Ray ORT.

Schwanger et al. (2005) predicted a rather high training effect in the PIT and a small influence of training on image-based factors which depend mainly on visual abilities and aptitudes. The significant interaction between test type and measurement is consistent with this assumption and shows that image-

based factors, i.e. the ability to cope with bag complexity, superposition and rotation of the threat item, can not be increased very much by training when compared to knowledge-based factors which depend highly on training.

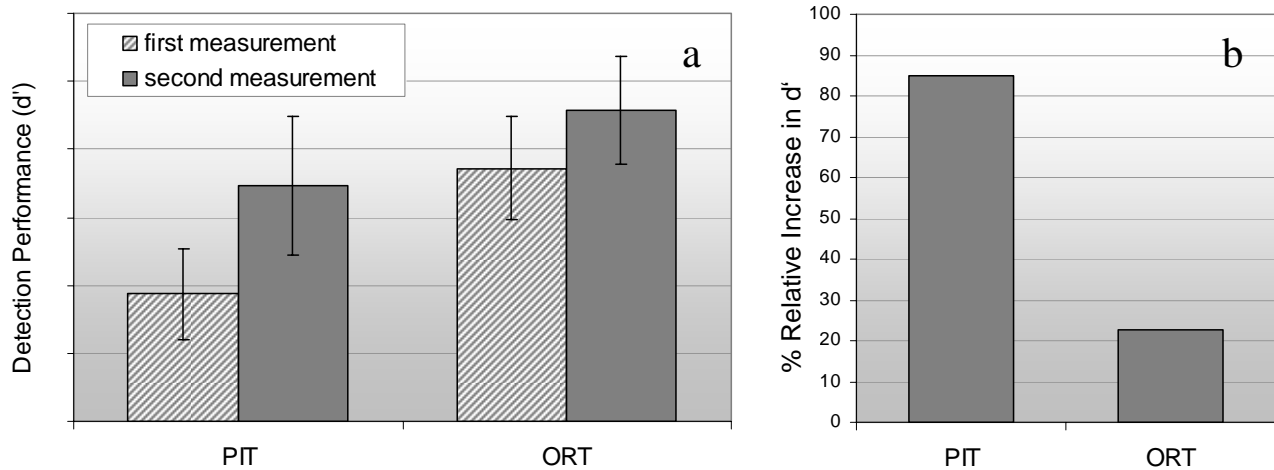


Figure 1: Detection Performance with standard deviations in the PIT and ORT for the first and second measurement left and percent difference in these two tests right.

Figure 1b shows the difference in percentage for both tests, taking the first measurement as baseline using the following formula (m1 means first measurement and m2 means second measurement):

$$\% \text{ Relative Increase in } d' = \left(\frac{m2 - m1}{m1} \right) * 100$$

As predicted by Schwaninger et al. (2005), there were large effects of training on knowledge-based factors (85.0%) and rather low influence of training on image-based factors (22.7%).

Discussion

The detection of threat items in passenger bags depends on knowledge-based and image-based factors. Screeners have to know which objects are prohibited and what they look like in x-ray images in order to recognize them (knowledge-based factors). In addition, they have to be able to cope with effects of bag complexity, superposition, and viewpoint (image-based factors).

In this study we investigated the role of training on knowledge-based and image-based factors in aviation security screening using the two x-ray screening tests, PIT and X-Ray ORT as well as XRT, a computer based individually adaptive training system. Overall, results show that the increase in detection performance after two years of recurrent computer-based training was much smaller in the X-Ray ORT compared to the training effect in the PIT. These results support the assumption that the PIT measures rather knowledge-based factors which can be increased remarkably through training compared to image-based factors measured by the X-Ray ORT, which are more difficult to increase by training. These results are as well consistent with results from a previous study by Schwaninger et al. (2005) which could show that the difference in x-ray detection performance between novices and aviation security screeners is much higher in the PIT than in the X-Ray ORT.

As can be seen in Figure 1a, detection performance in the PIT increased remarkably after two years of recurrent computer based training. Thus, an individually adaptive computer based training system like XRT helps to learn and store all kinds of prohibited items in different views in the visual memory. Additionally, it provides an excellent tool to react immediately to new threats as the training library of XRT can be updated constantly in an easy and fast way. However, besides the knowledge of a screener the abilities to cope with the image-based factors viewpoint, superposition and bag complexity are also very important. Hardmeier et al. (2005) revealed large inter-individual differences regarding the visual abilities

needed to cope with these image-based factors, which affected on the job performance. As this study shows, these abilities can only be trained to a limited extent and therefore should be measured in a pre-employment assessment to identify the candidates who are well-suited regarding these visual abilities. To summarize, both knowledge-based and image-based factors are very important in x-ray screening and can be measured relatively independent of each other using the PIT and X-Ray ORT. While knowledge-based factors can be enhanced remarkably by adaptive computer-based training, image-based factors should already be measured and used for selecting candidates in pre-employment assessment.

Acknowledgment

This research was financially supported by Zurich Airport Unique, Switzerland. We are thankful to Zurich State Police, Airport Division for their help in creating the stimuli and the good collaboration for conducting the study.

References

- Ghylin, K.M., Drury, C.G., & Schwaninger, A. (submitted). Two-component model of security inspection: application and findings. 16th World Congress of Ergonomics, IEA 2006, Maastricht, The Netherlands, July, 10-14, 2006.
- Green, D.M. & Sweets, J.A. (1966). *Signal Detection Theory and Psychophysics*, New York: Wiley.
- Hardmeier, D., Hofer, F., & Schwaninger, A. (2006). Increased detection performance in airport security screening using the X-Ray ORT as pre-employment assessment tool. Proceedings of the *2nd International Conference on Research in Air Transportation, ICRAT 2006*, Belgrade, Serbia and Montenegro, June 24-28, 393-397.
- Hardmeier, D., Hofer, F., & Schwaninger, A. (2005). The object recognition test (ORT) – a reliable tool for measuring visual abilities needed in x-ray screening. *IEEE ICCST Proceedings*, 39, 189-192.
- Hofer, F. & Schwaninger, A. (2004). Reliable and valid measures of threat detection performance in X-ray screening. *IEEE ICCST Proceedings*, 38, 303-308.
- Schwaninger, A., Hardmeier, D., & Hofer, F. (2005). Aviation security screeners visual abilities and visual knowledge measurement. *IEEE Aerospace and Electronic Systems*, 20(6), 29-35.
- Schwaninger, A. & Hofer, F. (2004). Evaluation of CBT for increasing threat detection performance in X-ray screening. In: K. Morgan and M. J. Spector, *The Internet Society 2004, Advances in Learning, Commerce and Security* (pp. 147-156). Wessex : WIT Press.
- Schwaninger, A. (2005b). Increasing Efficiency in Airport Security Screening. *WIT Transactions on the Built Environment*, 82, 405-416.
- Schwaninger, A. (2004). Computer based training: a powerful tool to the enhancement of human factors. *Aviation Security International*, FEB/2004, 31-36.
- Schwaninger, A. (2005a). Object recognition and signal detection. In: B. Kersten (Ed.), *Praxisfelder der Wahrnehmungspsychologie* (pp. 108-132). Bern: Huber.