RELATIONSHIP BETWEEN LEVEL OF DETECTION PERFORMANCE AND AMOUNT OF RECURRENT COMPUTER-BASED TRAINING

Stefan Michel, M Sc UZH
University of Zurich
Visual Cognition Research Group
Department of Psychology
Binzmuehlestrasse 14/22
8050 Zurich
Switzerland
s.michel@psychologie.uzh.ch

and

Max Planck Institute
for Biological Cybernetics
Dept. Bülthoff
Spemannstrasse 38
D-72076 Tübingen
Germany

Saskia M. Koller, M Sc UZH
University of Zurich
Visual Cognition Research Group
Department of Psychology
Binzmuehlestrasse 14/22
8050 Zurich
Switzerland
s.koller@psychologie.uzh.ch

Adrian Schwaninger, Prof. Dr.
University of Zurich
Visual Cognition Research Group
Department of Psychology
Binzmuehlestrasse 14/22
8050 Zurich
Switzerland
a.schwaninger@psychologie.uzh.ch

and

Max Planck Institute
for Biological Cybernetics
Dept. Bülthoff
Spemannstrasse 38
D-72076 Tübingen
Germany

Abstract – The usefulness of X-Ray Tutor (XRT) as an individually adaptive computer-based training system for aviation security screening officers to increase X-ray image interpretation competency has been proven in several studies. Many airports require their screening officers to conduct weekly recurrent training in order to enhance their capabilities and competencies. During training, X-ray images of passenger bags similar to how they appear at the security checkpoint have to be judged regarding the dangerousness of their content. Screening officers have to detect threat objects within the bags and discriminate harmless objects from threat objects. The advantage of XRT is its level-based construction. The training begins at the lowest difficulty level and increases in difficulty level with achieved performance for each screening officer individually. The aim of this article is to create a guideline and norm regarding training requirements in order to ensure that screening officers maximize their training benefit. Based on the X-Ray Tutor level algorithm we can express minimum requirements for screening officers in terms of image difficulty levels. For example, the minimum level to be achieved in X-Ray Tutor after one year of recurrent training was determined based on normative data assuming that screeners take about 1-2 training sessions of 20 minutes per week. Alternatively, it can also be expressed as requirement to reliably detect threat items depicted in difficult views with high superposition and medium bag complexity. Recommendations for standard reference levels and detailed background information on the X-Ray Tutor level are given.

Index Terms — Airport security, human factors, computer-based training, X-ray screening.

I. INTRODUCTION

The high potential for wide-scale fatalities on aircrafts shows the importance of scientific research in the field of airport security. Scientific research has gained high priority in the last years. Some X-ray machines provide automated detection of threat objects. Because of high false alarm rates of such technologies, the focus in X-ray image interpretation lies on the human operator who always makes the last decision if a passenger's bag is harmless or if it contains a threat object. One of the most important aspects therefore is that airport security screeners get individually adaptive training on X-ray image interpretation to enhance their knowledge about threat objects.

The X-Ray Tutor training system (XRT) was developed by the Visual Cognition Research Group (VICOREG) at the University of Zurich as a scientifically based and well-proven training program to enhance X-ray image interpretation competency very effectively ([1]; [2]). The training program displays X-ray images of passenger bags, where screeners have to visually inspect the images and search them for threat objects like they do at the security checkpoint. XRT is based on findings about how the human brain processes visual information in order to recognise objects in different views, when superimposed by other objects, and depending on bag complexity. One core advantage of XRT is that it is individually adaptive and level-based. In other words, when screeners train with X-Ray Tutor, they reach higher difficulty levels based on their individual detection performances. The question arises which level is recommended to be reached after a specific amount of training time to guarantee a reliable and effective recognition of forbidden objects at the security checkpoint. Therefore, one important aspect is that a standard for detection performance improvement is needed, which is the main goal of this article.

Two approaches should be pursued:

1) A standard should be defined by taking into account view difficulty, superposition, and bag complexity through XRT levels.

2) The second approach is based on the data of several airports. In the course of a project, aviation security screeners at these airports conducted recurrent computer-based training with XRT during at least one year. We consider proposing a standard regarding which level in XRT screeners must have reached if they used the system for 12 months for 20 minutes per week on average.[1]

II. IMAGE-BASED AND KNOWLEDGE-BASED FACTORS

Humans are adept at detecting different objects without any problems in a very short time as far as conditions are favourable. As soon as the conditions become unfavourable, detection performance can decrease. For example, if an object is superimposed by other objects, its shape is hard to separate (e.g., figure-ground segregation) and therefore recognition can become difficult. In the same manner, recognition of rotated objects can be difficult when objects are seen from an unusual perspective. An object can only be recognised when this particular object or a similar looking one has been seen before and been stored in visual memory. However, if there is a large difference in the appearance (e.g., because it is seen from a different angle), it cannot be recognised well anymore (always under the assumption that it has never been seen before). An object has to be seen from different angles in order to store all these views in visual memory. Then, the object can be recognised no matter which angle it is seen in. Studies regarding object recognition indicate that for most objects six views are sufficient to capture the qualitative differences resulting from viewpoint changes because the human brain is able to interpolate between stored views ([1]). Additionally, if a scene (or a bag for instance) is very complex, problems in recognizing individual objects can occur. Too much information distracts attention and impedes detection and recognition of objects. These limitations that derive from the image itself are defined as image-based factors. [3] defined three image-based factors as important for X-ray image interpretation (see Figure 1): It is harder to detect an object in a rotated view compared to the upright view (effect of viewpoint). The superposition of an object by others can impair the detection performance as well (effect of superposition). A bag containing different objects and the type and number of these objects can distract visual attention and therefore also impair the detection performance (effect of bag complexity). These image-based factors (view, superposition, and bag complexity) should be taken into account in an individually adaptive training for aviation security screeners.

In other words, an individually adaptive training system should increase the difficulty of the training material (i.e., X-ray images of passenger bags, some of which contain threat objects) regarding these image-based factors by means of the individual performances. For example, people are first trained with objects in easy rotations. If a certain level of detection performance in reached, the level is increased and more difficult views of threat objects are displayed. In higher levels, threat objects are more superimposed by other objects. Finally bag complexity is increased; individually adapted to the performance and difficulties a screener has in coping with each of these image-based factors. If, for example, a screener has problems coping with rotated threat objects but not when threat objects are superimposed by other objects, then the screener gets augmented training on different viewpoints of threat objects. The training proceeds in this way until the screener is able to detect rotated threat objects reliably and the screener succeeded in overcoming the difficulties in that detection performance increases. In addition, not only image-based factors are trained but also knowledge-based factors. Knowledge-based factors refer to the knowledge about which objects are prohibited and how they look like in X-ray images of passenger bags. To improve the visual knowledge, a large and representative image library is required.

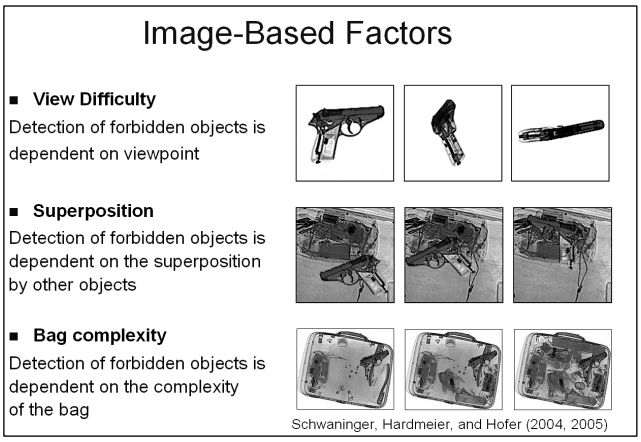

Fig. 1: Explanation of image-based factors

III. X-RAY TUTOR AS AN INDIVIDUALLY ADAPTIVE TRAINING SYSTEM

X-Ray Tutor (XRT) is a scientifically based training program in which screeners have to decide if an X-ray image of a passenger bag is harmless or not (see Figure 2). The training is individually adaptive, that is to say, it automatically adapts to the performance of individual airport security screeners considering the difficulty of the images. X-Ray Tutor automatically combines images of fictional threat items with X-ray images of passenger bags. This is performed by an individually adaptive algorithm. X-Ray Tutor contains a large image library of threat objects that are X-rayed from different standardized views. Most of the objects can be depicted from up to 72 different viewpoints, which allows training screeners to detect threat objects independent of viewpoint. Additionally, threat objects from different threat categories (e.g., guns, knives, IEDs, other prohibited items) are integrated in XRT to make sure that a screener will be able to detect a large number of different threats. This image library was built in close collaboration with experts of Zurich State Police Airport Division and other organizations, and it is being extended continuously. The large image library is used in the individually adaptive training system in a way that objects which are poorly

[1] This research was financially supported by the European Commission Leonardo da Vinci Programme (VIA Project, DE/06/C/F/TH-80403) and the Ministry of Justice, The Netherlands.

recognized are presented more often to a screener. How superimposed such a threat object is or the complexity of a bag that threat objects is presented in is depending on the difficulty level (see Table I for an overview).

TABLE I
DIFFICULTY LEVELS IN X-RAY TUTOR

Level	Viewpoint	Super-position	Bag Complexity
1	Easy	Low	Low
2	Difficult	Low	Low
3	Easy	High	Low
4	Difficult	High	Low
5	Easy	Low	Middle
6	Difficult	Low	Middle
7	Easy	High	Middle
8	Difficult	High	Middle
9	Easy	Low	High
10	Difficult	Low	High
11	Easy	High	High
12	Difficult	High	High

How fast a screener reaches a higher level in XRT is dependant on the number of threat objects in the image library and on the screener's performance. X-Ray Tutor is available in two different versions. XRT standard edition contains 100 threat objects in up to 72 views each. XRT professional edition contains 400 threat objects also in up to 72 views each. In this article, the results are based on training with XRT standard edition.

Fig. 2: Screenshot of the XRT CBS 2.0 training system during training. At the bottom-right feedback is provided after each response. If a bag contains a prohibited item, an information window can be displayed (see bottom left of the screen).

During training with X-Ray Tutor, X-ray images of bags are presented on the screen for 15 seconds (standard setting). Screeners have to decide whether the bag is OK (i.e. it contains no threat object) or whether it is NOT OK (i.e. it contains a threat object). After each response, feedback is provided informing the screener whether his/her response was correct. If the bag contained a threat object the screener can view detailed information and a real image of the threat object (see Figure 2). For further information on X-Ray Tutor see [1].

IV. PREVIOUS STUDIES SUPPORTING THE EFFECTIVENESS OF X-RAY TUTOR

Different studies could show that there are large effects of viewpoint, superposition, and bag complexity due to training with XRT ([3]; [4]; [5]). Supporting these findings, a recent study by [6] shows that there is an increase of detection performance not only for objects in easy but also for objects seen in difficult views.

A study by [2] shows that there is a very high increase of detection performance with the individually adaptive training system XRT compared to a training system which is not individually adaptive and uses a smaller image library (see Figure 3). These results are based on two measurements with the same test (X-Ray CAT, see [7]) with 6 months of X-ray Tutor training between these measurements.

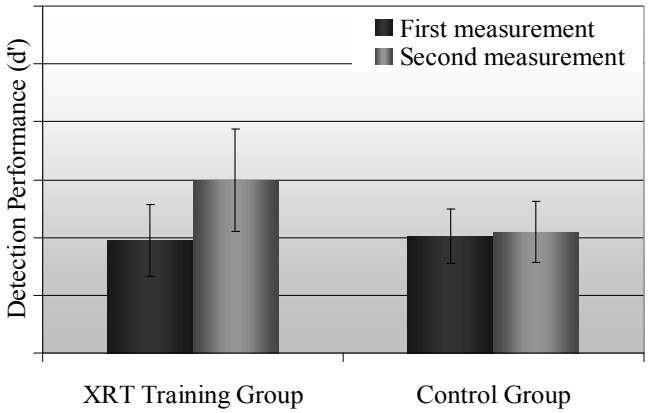

Fig. 3: Detection performance with standard deviations for the XRT training group vs. the control group comparing first and second measurement.

Another important finding is that not only trained threat objects can be recognized better due to training. It is also possible to increase recognition of visually similar objects (transfer effect). A study conducted recently by [2] supports this finding. Therefore, during training the knowledge about how X-ray images of threat objects look like transfers to the detection of other similar looking objects. These findings are very important considering that not each threat object existing in the world has to be included in training. It is necessary to include objects from various categories and to have a representative threat image library. All in all the effectiveness of XRT is well proved in different studies.

V. METHODS

We will mainly present the progress of the screeners' training program regarding their difficulty level increase compared to the hours of training. XRT has been operational at different airports for the same duration. Screeners were obligated to use XRT for 20 minutes per week for at least 12 months. For example, if a screener has used the training system once a week for 20 minutes during 12 months, he has used the system for a total of 16 hours.

In total, data of 381 aviation security officers from four airports is included in this report. They trained for 14.69 hours on average with a standard deviation of about 12.97 hours during 12 months. For an overview of each of the four airports see Table II. In this article the results are based on XRT standard edition which contains 100 threat objects in up to 72 views.

TABLE II
NUMBER OF SCREENERS AND TRAINING HOURS FOR EACH AIRPORT

Airport	No. of Screeners	Training in months	Training hours (average)	Training hours (stdev.)
Airport 1	26	12	18.96	10.65
Airport 2	83	12	30.55	6.83
Airport 3	202	12	7.05	8.40
Airport 4	70	12	16.33	11.88

VI. RECOMMENDATION FOR DETECTION PERFORMANCE IMPROVEMENT

For the recommendations, two approaches should be pursued:

1) A standard should be defined by taking into account view difficulty, superposition, and bag complexity through XRT levels (see Table I).

2) A standard should be defined by taking into account training duration and level progress.

A. Recommendation based on object recognition theories

A standard should be defined by taking into account view difficulty, superposition, and bag complexity through XRT levels. This minimum standard is based on object recognition theories which imply that an object can be recognised best when it has already been seen from different viewpoints. Additionally, it is also important that the forbidden objects are to some extent superimposed by other objects in the training images to make sure that objects can be well recognised independent of the position inside the bag. Finally, it is harder to find an object in a close-packed bag. Therefore, screeners should be able to recognise a threat object even if the bag is fully packed with many different harmless objects.

X-Ray Tutor contains 12 levels regarding these three image-based factors (see Chapter III). In the first level, easy viewpoints of threat items with low superposition and low bag complexity are presented. The second level contains difficult viewpoints with low superposition and low bag complexity. In level three the superposition is increased whereas the viewpoint of threat items is easy and bag complexity is low. See Table I for the combination of all 12 levels in X-Ray Tutor according viewpoint, superposition, and bag complexity.

Regarding these image-based factors a screener should reach level 6 after one year of training to fulfil the theoretical assumptions from object recognition theories.

In level 6, a screener has seen the threat objects from easy and difficulty viewpoints. Furthermore, threat objects depicted in easy as well as in difficult orientations are presented with low and high superposition, respectively. In addition, low and middle bag complexity levels have also been seen when people have reached level 6. Therefore, screeners are exposed to almost all different image-based factors except high bag complexity. This recommendation guarantees a reliable recognition for threat objects in almost all cases at the security checkpoint. Only very high complexities are not well trained so far. When considering a bag too complex, screeners at the security checkpoint are obligated to search such a bag manually.

B. Recommendation based on data of airports which have trained for 12 months

A very high correlation between the amount of training hours and the XRT level achievement was found with $r = .93$.

Figure 4 shows the XRT level increase as a function of training duration (means and standard deviations) from real data of 381 screeners from four different airports. Note that

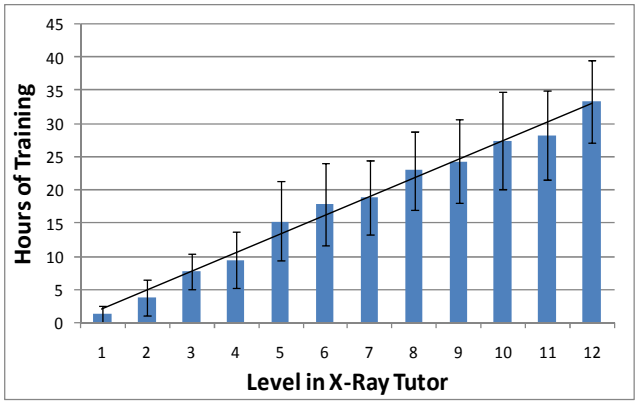

Fig. 4: Average of training hours per level

the variation among screeners is large, with regards to how long it takes to reach a certain training level. For example, to reach level 2, 4 hours of training are needed in average. The standard deviation represents a substantial difference between people in making progress in levels.

In Figure 5, the percentage of screeners in a specific level is shown. Most screeners (25.46%) are in level 1, but none of them trained the requested amount of time of around 20 minutes per week due to different reasons (e.g., job fluctuation). Screeners in level 1 trained for about 1.37 hours within 12 months. Screeners who used the system as requested (i.e. at least 20 minutes per week during 12 months) reached level 6 after one year of training.

To conclude, based on these analyses from real data the recommendations for screeners is to reach level 6 after one year of training when XRT is used for about 20 minutes per week (in total 16 hours of training). 61.68% of all screeners in this report fulfilled this requirement and reached level 6 or

higher after 12 months of training.

In addition, Figure 6 shows the minimum amount of training hours needed to reach a specific level based on the same actual data from 381 screeners. For example, the minimum amount of training time required to reach level 6 is 9.22 hours. To reach level 12 the best screener needed 21.88 hours of training.

During a training session of about 20 minutes, a screener sees about 157 X-ray images on average. In Figure 7, the numbers of images that have been seen during training are shown for each level in XRT. A very large correlation between XRT level achievement and number of images that have been seen during training was found with r = .96. The same large correlation between training duration and number of images was found with r = .96. In Figure 7 it can be seen that screeners have seen about 8'094 images on average until they reach level 6 and more than 16'606 X-ray images on average to reach level 12. With such a large number of images screeners are well exposed and familiarized to thousands of different X-ray images of passenger bags.

Finally, if a screener does not reach level 6 after 12 months of training, it is important that the screener is using the training system more regularly and more frequently. When screeners have reached level 12 it is important to update the image library with new threat objects for further increasing the knowledge about threat objects.

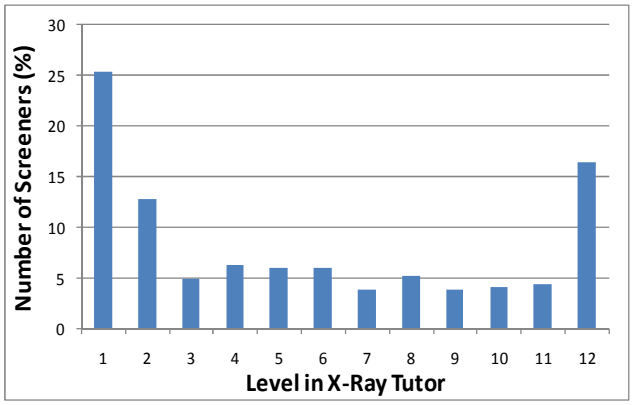

Fig. 5: Number of screeners (%) per level

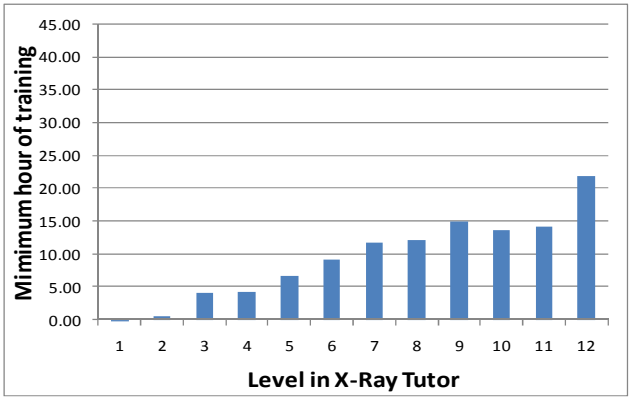

Fig. 6: Minimum amount of training hours to reach a specific level in XRT.

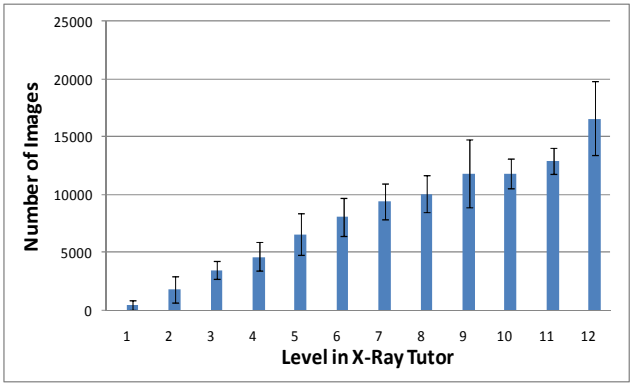

Fig. 7: Number of X-ray images per level (average)

VII. CONCLUSIONS

Because the human operator always has the last decision at the security checkpoint, it is important to point out that training is one of the core aspects in aviation security to guarantee a reliable and effective recognition of threat objects in passenger bags. Without training, a screener is not able to detect threat objects reliably, especially because some threat objects (e.g., IEDs) are seen very rarely at the security checkpoint. A training system should be individually adaptive to train screeners optimally based on their performances. The effectiveness of XRT has been shown in many different studies ([4]; [6]; [8]; [9]; [10]). For example, the study by [2] showed very clear benefits for XRT. In this study, a control group trains with a non-individually adaptive training system and is compared to a group which used an individually adaptive training system (X-Ray Tutor) for the same period. Large training effects could only be found for the individually adaptive training group and not for the control group (see Figure 3). With all these findings, the importance of individually adaptive training cannot be neglected.

Another important aspect of the training system is that threat objects are presented in different difficulties like easy and difficult orientations, low and high superposition, and low and high bag complexity. This is represented in different difficulty levels implemented in XRT. The progress in these levels is dependent on the individual threat detection performances.

The aim of this article is to create a guideline and norm regarding training requirements in order to ensure that screening officers maximize their training benefits. Therefore, two recommendations based on object recognition theories and based on real XRT training data, respectively, are taken into account.

Taking the findings together, the minimum standard after 20 minutes of training per week during 12 months (is equivalent to 16 hours of training in total) is that screeners have to reach level 6 in XRT. When screeners have reached level 6, threat objects have been seen from different viewpoints, in low and high superposition, and also in medium bag complexities. This fulfils the theoretical assumptions from object recognition theories for a reliable recognition of objects.

In addition, real data shows that, on average, screeners reached level 6 when they used XRT for 12 months for 20 minutes per week. This supports the theoretical assumption and shows that it is realistically possible to fulfil these requirements.

Finally, if XRT standard edition is used to prepare screeners for the initial certification tests, it is recommended to train with XRT standard edition for about 16 hours at least or to reach level 6. Then, screeners should be well prepared for the certification test and for working at the security checkpoint.

To guarantee that the threat detection performance is stable, recurrent training with an appropriate training system and training duration (minimum 20 minutes per week) is recommended.

VIII. REFERENCES

[1] A. Schwaninger, "Computer based training: a powerful tool to the enhancement of human factors". *Aviation Security International*, FEB/2004, pp 31-36.

[2] S.M. Koller, D. Hardmeier, S. Michel, & A. Schwaninger. „Investigating training and transfer effects resulting from recurrent CBT of x-ray image interpretation". In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society*. Austin, TX: Cognitive Science Society, 2007, pp 1181-1186.

[3] A. Schwaninger, D. Hardmeier, & F. Hofer. „Aviation security screeners visual abilities & visual knowledge measurement". *IEEE Aerospace and Electronic Systems, 20(6), 2005, pp* 29-35.

[4] D. Hardmeier, F. Hofer, & A. Schwaninger. „The role of recurrent CBT for increasing aviation security screeners' visual knowledge and abilities needed in x-ray screening". *Proceedings of the 4th International Aviation Security Technology Symposium, Washington, D.C., USA, November 27 – December 1, 2006, pp* 338-342.

[5] F. Hofer, D. Hardmeier, & A. Schwaninger. „Increasing airport security using the X-ray ORT as effective pre-employment assessment tool". *Proceedings of the 4th International Aviation Security Technology Symposium, Washington, D.C., USA, November 27 – December 1, 2006,* pp 303-308.

[6] S. Michel, J.C. de Ruiter, M. Hogervorst, S.M. Koller, R. Moerland, & A. Schwaninger. „Computer-based training increases efficiency in x-ray image interpretation by aviation security screeners". *Proceedings of the 41st Carnahan Conference on Security Technology,* Ottawa, October 8-11, 2007.

[7] S. Koller, & A. Schwaninger. „Assessing X-ray image interpretation competency of airport security screeners". Proceedings of the *2nd International Conference on Research in Air Transportation, ICRAT 2006,* Belgrade, Serbia and Montenegro, June 24-28, 2006, pp 399-402.

[8] A. Schwaninger, & F. Hofer. „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*, Wessex: WIT Press, 2004. pp 147-156

[9] A. Schwaninger, F. Hofer, & O.E. Wetter. „Adaptive computer-based training increases on the job performance of x-ray screeners". *Proceedings of the 41st Carnahan Conference on Security Technology,* Ottawa, October 8-11, 2007.

[10] S.M. Koller, D. Hardmeier, S. Michel, & A. Schwaninger. „Investigating training, transfer, and viewpoint effects resulting from recurrent CBT of x-ray image interpretation". *Journal of Transportation Security, 1(2)*, 2008, pp 81-106.

IX. VITA

Stefan Michel is a doctoral student at the Max Planck Institute for Biological Cybernetics in Tuebingen (Germany) and at the University of Zurich (Switzerland). He is also project manager for several aviation security training studies. He graduated in psychology, child and adolescent psychiatry and neuropsychology at the University of Zurich. His main research topics are perceptual learning, psychophysics, signal detection and applied vision research.

Saskia M. Koller is a doctoral student at the Visual Cognition Research Group of Prof. Dr. Adrian Schwaninger at University of Zurich. In 2006 she finished her studies in Psychology, Business Management and Criminology at University of Zurich and is now writing her doctoral thesis in the field of airport security.

Prof. Dr. Adrian Schwaninger has been lecturing at the University of Zurich and at the Federal Institute of Technology (ETH) in Zurich since 1999 and at the University of Applied Sciences Northwestern Switzerland since 2008. He is a member of the Training and Technical Task Forces of the European Civil Aviation Conference (ECAC), the moderator of the ECAC Technical Task Force TIP Study Group, and the Chairman of the InterTAG ad hoc Working Group on Competency Assessment. In 1999 he received the Young Researcher Award in Psychology. In 2003 he received the ASI International Award of Excellence in Aviation Security: Enhancement of Human Factors. Together with his Visual Cognition Research Group (VICOREG) he is in charge of several aviation security projects in Belgium, Bulgaria, Canada, France, Germany, Greece, Norway, Romania, Sweden, Switzerland, The Netherlands, and the United States of America.