SCIENCE & SECURITY March 2017

CASRA NEWSLETTER – ISSUE 15

This spring 2017 issue features two interesting new topics. In the section "Research Put Across", we present a study on assisted target recognition algorithms and training for improving detection performance of customs officers. In the section "Security in Practice", we examine whether knowledge on everyday objects is related to effectiveness and efficiency in X-ray image inspection performance.

We hope that you enjoy reading this issue, and we welcome any feedback you might have as well as topics you would like us to address in upcoming newsletters.

We wish you a nice spring season,

Dr. Diana Hardmeier
Director

Prof. Dr. Adrian Schwaninger
Chairman

TOPICS IN THIS ISSUE:

RESEARCH PUT ACROSS

HOW USEFUL ARE ASSISTED TARGET RECOGNITION ALGORITHMS AND TRAINING FOR IMPROVING THE DETECTION PERFORMANCE OF CUSTOMS OFFICERS?

In the CASRA Newsletter one year ago (Issue 12, March 2016), we have described ACXIS (Automated Comparison of X-ray Images for Cargo Scanning). This is a research project funded by the European Union under the 7th framework program with the main aim to increase effectiveness and efficiency of cargo screening at customs. Among other tasks, the project researched and developed assisted target recognition (ATR) algorithms to assist customs officers in the challenging search for illicit goods in cargo containers. With the final decision of whether a container should be opened for further investigation still in the responsibility of the officers, the detection performance of the ATRs is not the only determinant of their benefit. Also the successful interaction between the officers and the ATRs is crucial for the joint performance. This joint performance was investigated in a validation study, also considering the importance of training for detection performance.

SECURITY IN PRACTICE

IS EVERYDAY OBJECT KNOWLEDGE RELATED TO X-RAY SCREENERS' EFFECTIVENESS AND EFFICIENCY?

The main objective of CASRA is to increase security and facilitation at airports and other environments involving people and technology, using an adaptive approach by combining applied psychology and computer science. Learning which objects are prohibited and what they look like in X-ray images of passenger bags is important for X-ray screeners to achieve a good detection in visual inspection of passenger bags. However, as these bags also contain a large variety of harmless everyday objects, it could be assumed that knowing how such objects look like in X-ray images might help screeners to work better. CASRA researchers conducted a scientific study to investigate whether the knowledge of everyday objects is related to X-ray screeners' effectiveness and efficiency (Experiment 1) and if novices' everyday object knowledge can be trained (Experiment 2).

HOW USEFUL ARE ASSISTED TARGET RECOGNITION ALGORITHMS AND TRAINING FOR IMPROVING THE DETECTION PERFORMANCE OF CUSTOMS OFFICERS?

Text: Yanik Sterchi

Automated Comparison of X-ray Images for Cargo Scanning (ACXIS) is a research project funded by the European Union under the 7th Framework Program (see www.acxis.eu for further information). The ACXIS consortium (including CASRA) conducts research and development for improving effectiveness and efficiency of cargo screening. As part of this, ACXIS partners are developing Assisted Target Recognition (ATR) methods to automatically detect potentially illicit goods[1] by means of dedicated algorithms. For this purpose, first the Dutch Customs Organization (DTCA) and the Swiss Federal Customs Administration (FCA) defined several illicit goods detection scenarios. The consortium then prioritized and selected scenarios, resulting in four ATRs being developed with Smiths Detection in the lead. As an example, methods for cigarette detection based on texture analysis have been developed and show excellent first results. While the overall project has been described in the CASRA Newsletter issue 12 (March 2016), the present article is focusing on the validation study of the ACXIS project.

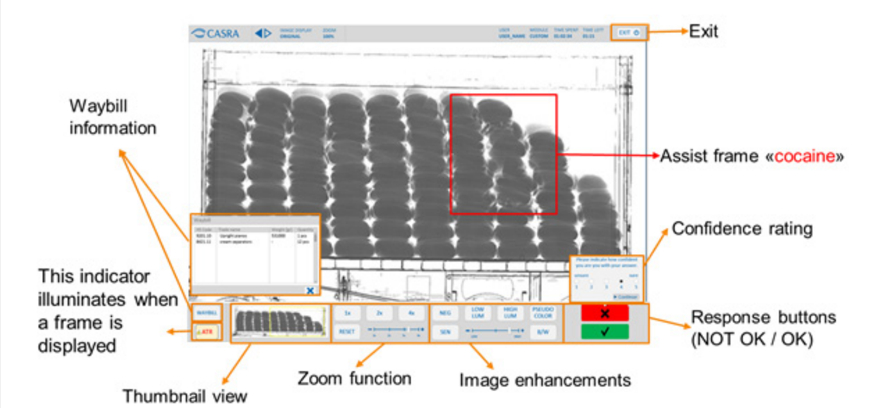

Figure 1: Simulator interface for training and testing with ATRs.

How effective ATRs are, does not only depend on the performance of the automatic algorithms. It is also important, how the information provided by the ATRs is implemented into the screening process. The idea within ACXIS was that ATRs support visual inspection of customs officers by directing their attention on suspicious areas in the X-ray image and help decision making (red frame in Figure 1). In this implementation scenario, the screening officer still makes the final call on whether the cargo needs further inspection. Therefore, whether and how much the detection performance of the screener improves due to ATRs, depends on a variety of factors. Research on human–machine interaction in various professions revealed that the benefit of automated support systems is influenced by the officers' experience with the system ([1]), their expertise in the task ([2]), and the occurrence of hits and false alarms[2] generated by the system (e.g. [1], [3]). Research has also shown that automated support systems do not always lead to better performance (e.g. [4]). A key component of the ACXIS project was therefore to evaluate the designed implementation of the ATRs. Since the officers' experience with the system and their expertise in the task can influence the benefit of an automated system, the successful use of the ATRs might depend on whether screeners are trained and had the chance to gain experience with the ATRs. Computer-based training has repeatedly been shown to increase detection performance in X-ray image inspection of air passenger bags (e.g. [5], [6]) and one study also showed its benefit in cargo screening ([7]). The ACXIS validation study therefore also aimed to investigate the influence of training. The validation study was designed with following goals:

- Replication of the previously found effects of computer-based training on detection performance in cargo screening
- Evaluate whether ATRs lead to an increase in detection performance
- Investigate whether the benefit of the ATRs is influenced by the computer-based training

METHOD: SIMULATOR

For the validation study, CASRA has developed a new platform for cargo X-ray image interpretation training and testing, called the Customs X-Ray Simulator (described in the CASRA Newsletter issue 13, July 2016).

ATRs

In the validation study, four assist functions have been evaluated. ATRs 1-3 aimed at detecting specific potentially illicit goods within the freight: hidden cigarettes (ATR1), weapons (ATR2) and narcotics (ATR3). A fourth ATR detected anomalies in the container structure (ATR4), and was therefore able to detect any sort of potentially illicit

[1] While some goods are generally forbidden (e.g. cocaine), many goods require permits and are illicit if not properly declared on the waybill.
[2] A hit is defined as a correct alarm by the system, e.g. if a potentially illicit good is hidden in the cargo and the system generates an alarm. If an alarm incorrectly indicates the presence of a potentially illicit good while no such good is hidden in the cargo, this is considered a false alarm.

goods, but only in the container structure and not within the freight (Figure 2). For a better control over the experimental conditions, the officers were trained and tested separately for the location-specific ATR4 and the target-specific ATRs 1-3. The detection algorithms of the ATRs have been still under development when the validation study was conducted, and the alarms of the ATRs were therefore set manually by screening experts. All four ATRs were designed to detect 75% of their respective targets, while 25% of the X-ray images produced a false alarm.

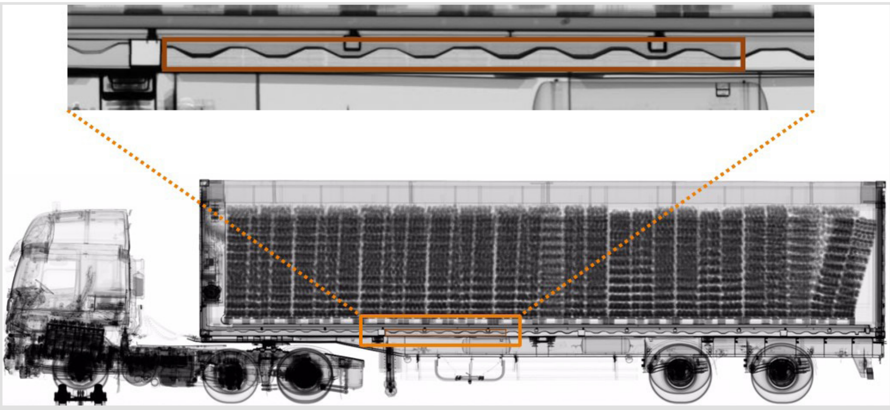

Figure 2: X-ray image of a container with hidden illicit goods.

Content

To create the training and testing material, the Dutch Customs Organization (DTCA) provided images of cargo recorded with a HCV scanner (also appropriate waybills were included but no top views of the X-ray scans). Empa and DTCA recorded cigarettes, weapons, and mockup narcotics, which were merged into cargo images by CASRA, using a merging tool from CEA. Officers from DTCA and Swiss Federal Customs Administration (FCA) reviewed the images for quality (these officers were excluded from the study).

PARTICIPANTS & DESIGN

46 employees of the DTCA and 21 employees of the FCA completed the study. All participants conducted the Ishihara test in order to confirm their color perception ability, and their object recognition skills were assessed with the X-Ray ORT ([8]). Participants that passed the Ishihara test were divided into three groups with comparable object recognition skills (with the exception of the control group of DTCA, which consisted of administrative personnel). As illustrated in Figure 3, the first group did not receive training or assistance by ATRs and served as control group. The second group (training group) was tested and trained without ATRs, while the third group was tested and trained with ATRs.Tests to measure detection performance in cargo screening took place once before training and then again at the end of the study. Each test consisted of two parts, part one with illicit goods only hidden within the freight and part two with half the illicit goods hidden in the container structure (while the other half remained hidden within the freight). In the ATR group (AG), part one used ATRs 1-3 and part two ATR4. Each part consisted of 96 images with illicit goods in one fourth of the images, resulting in eight targets for each of the three categories of illicit goods. For the defined hit and false alarm rate of the ATRs this meant that an alarm indicated an illicit good in 50% of the cases (and missing 25% of the illicit goods). Participants agreed to two training sessions of 15 minutes per week over the course of four months. The training consisted of three difficulty levels where participants progressed to the next level after three to four hours of training.

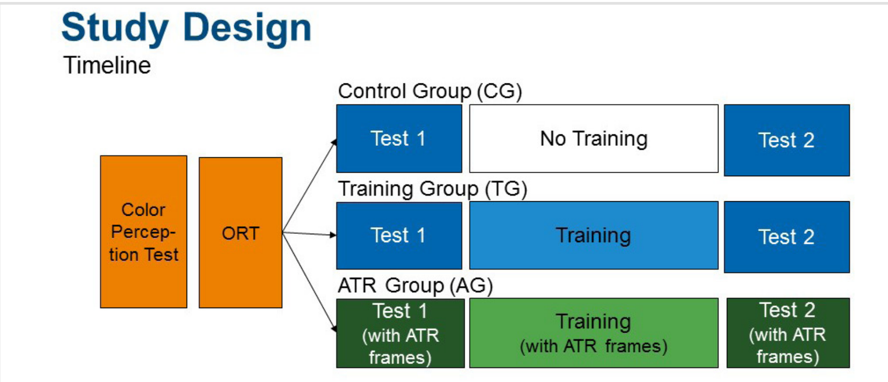

Figure 3: Experimental design of the study.

RESULTS AND DISCUSSION

Since the tests consisted of images from DTCA only, which differed notably from the images familiar to officers from FCA (resulting in higher false alarm rates), results were analyzed for DTCA and FCA separately and should not be compared between the two customs organizations.

Figure 4 summarizes the results of test part one (illicit goods within the

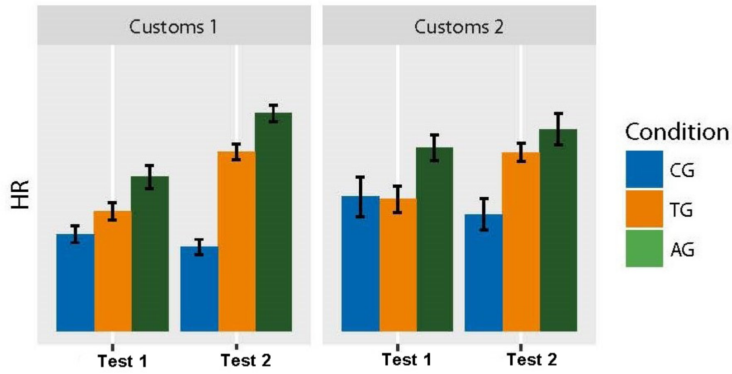

Figure 4: Mean hit rate of test part one of control group (CG, no training, no assistance by ATRs), training group (TG, training but no assistance by ATR), and ATR group (AG, training and assistance by ATRs). Test 1 was before the training for TG and AG, Test 2 afterwards. Error bars indicate standard errors of the mean.

freight only). As can be seen, for the training group detection (hit rate) was higher in Test 2 compared to Test 1, while for the control group, detection remained constant. Statistical analyses confirmed that training significantly improved detection for officers from both customs organizations. Further analyses revealed that this improvement did not differ significantly between the three categories of illicit goods and that the false alarm rate was not affected by training. Figure 4 also shows that detection was higher for the ATR group than the training group, meaning that ATRs for the detection of cigarettes, drugs, and guns could improve detection when looking at these categories combined. Analyses confirmed this improvement to be statistically significant and not to differ significantly between Test 1 and Test 2. This means that the benefit of the ATRs (more specifically: the difference between the training group and ATR group) was about the same before and after training. Further analyses revealed that the false alarm rate was not influenced by ATRs but that the benefit of ATRs differed between the three categories of illicit goods, being larger for the detection of guns than for the detection of cigarettes and narcotics. This might be explained by the size of the guns, which were relatively small and therefore easily missed if not highlighted by the ATR.

In test part two, where half of the illicit goods were hidden in the container structure (Figure 2) and the other half within the freight, training also had a positive effect on the hit rate without affecting the false alarm rate. But ATR4 (detection of anomalies in the container structure) showed both, benefits and disadvantages: the participants who were assisted by ATR4 detected more illicit goods that were hidden in the container structure, but missed more of the other illicit goods within the freight. A possible explanation for this might be that the ATR directed the officers' attention more towards the container structure and away from the freight.

Overall, the study replicated that training can be an important contribution to improve detection performance in cargo screening. The ATRs led to improved detection performance, but their benefit depended on the category of illicit goods, and ATR4 also showed some undesirable side effects. These results highlight the importance of research in human–machine interaction, before automated support systems are implemented in practice.

JOINT EXPLOITATION EVENT

On the 19th of May, 2017, the Joint Exploitation Event will take place in Paris. It provides the opportunity for technology providers of five European research projects, one of which is AXCIS, to present their exploitable project results to customs organizations in order to prepare the uptake of research results. For more information on the event or registration, please contact info@casra.ch.

REFERENCES

[1] Bliss, J., Dunn, M., & Fuller, B.S. (1995). Reversal of the cry - wolf effect: An investigation of two methods to increase alarm response rates. *Perceptual and Motor Skills*, *80*(3),1231–1242.

[2] Strauch, B. (2016). The Automation-by-Expertise-by-Training Interaction. Why Automation-Related Accidents Continue to Occur in Sociotechnical Systems. Human Factors: *The Journal of the Human Factors and Ergonomics Society.*

[3] Parasuraman R., Sheridan T.B., & Wickens C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans. 30*(3), 286–297.

[4] Hättenschwiler, N., Mendes, M., Sterchi, S., & Schwaninger, A. (2017). *Automation in airport security X-ray screening of cabin baggage: Examining benefits and possible implementations of automated explosives detection.* Manuscript submitted for publication.

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

[6] Halbherr, T., Schwaninger, A., Budgell, G.R., & Wales, A. (2013). Airport Security Screener Competency: A Cross-Sectional and Longitudinal Analysis. *Int. J. Aviat. Psychol.*, *23*(2), 113–129.
[Download PDF]

[7] Michel, S., Mendes, M., de Ruiter, J., Koomen, G., & Schwaninger, A. (2014). Increasing X-ray image interpretation competency of cargo security screeners. *International Journal of Industrial Ergonomics*, 44, 551-560.
[Download PDF]

[8] Hardmeier, D., Hofer, F., & Schwaninger, A. (2006a). 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, 2006,* 3-397.
[Download PDF]

IS EVERYDAY OBJECT KNOWLEDGE RELATED TO X-RAY SCREENERS' EFFECTIVENESS AND EFFICIENCY?

Text: Nicole Hättenschwiler

The main task of an X-ray screener is to visually inspect X-ray images of passenger bags (Figure 1) and to decide whether a bag is harmless or whether it might contain a prohibited item and therefore needs further inspection through secondary search. In initial classroom sessions, computer-based e-learning courses and on-the-job training, X-ray screeners learn how to interpret X-ray images in order to recognize everyday objects and prohibited items. In the last decade, several studies showed that computer-based training (CBT) is important to achieve good detection performance in X-ray image interpretation (for more information see [1-8]).

While these studies provided converging evidence for the importance to learn which items are prohibited and what they look like in X-ray images, the role of everyday object knowledge has not yet been addressed specifically. This is an interesting topic, especially from an operational point of view. An example: when looking at Figure 1, can you name all the everyday objects within this passenger bag? Can you tell for certain if the object behind the glasses is either a pen or a small knife? Now, if you are not completely sure while working at the checkpoint, you would probably send the bag to secondary search, right? Therefore, it may be assumed that the knowledge on what everyday objects look like in X-ray images could result in fewer cases where an everyday object is confused with a prohibited item (e.g. a pen can resemble a switchblade knife as shown in Figure 2). This would result in fewer false alarms, i.e. wrongly judging a bag to contain a prohibited item. False alarms have to be resolved through secondary search, which typically involves manual search and/or alarm resolution using explosive trace detection technology ([9]). Due to the additional time needed for secondary search, high false alarm rates can have a strong negative impact on throughput ([9]) and they may also result in lower passenger satisfaction ([10]). Therefore, it is worth investigating the role of everyday object knowledge and training because it could be related to efficient X-ray screening.

Figure 1: Screenshot showing an X-ray image of a passenger bag.

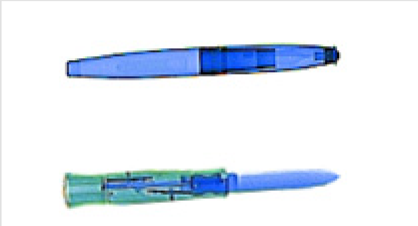

Figure 2: Screenshot showing X-ray images of a pen and a switchblade knife.

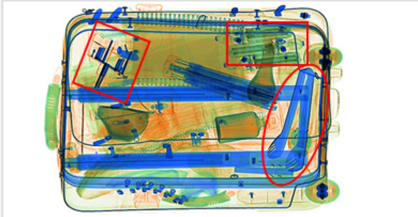

Figure 3: Screenshot showing an X-ray image of a passenger bag from the X-ray OCNT with framed objects.

IS EVERYDAY OBJECT KNOWLEDGE RELATED TO LOWER FALSE ALARM RATES IN X-RAY SCREENING TASKS?

To address this question, 15 X-ray screeners were first tested using an X-ray object categorization and naming test (X-Ray OCNT). For the X-Ray OCNT, the screeners had to categorize and name three objects in each passenger bag (previously marked by a red frame) as a threat-item or a harmless everyday object (Figure 3). Almost all screeners were able to correctly categorize the everyday objects (96%). When they were asked to name these objects, they got the answer right in 82% of the cases, which is still very good. The question now is whether these results mean that everyday objects do not need to be trained? The answer is "No" since there were large differences between screeners; performance varied from 54% to 96% correct naming.

In a second phase, the screeners analyzed images using a simulated X-ray baggage screening test (XBST). The XBST resembles the task at the checkpoint as the screeners had to judge several X-ray images of passenger bags and decide whether it contained a prohibited item (NOT OK) or not (OK). When screeners wrongly judged a bag to contain a prohibited item, it counted as a false alarm[1]. The false alarm rate was calculated as the number of false alarms divided by the number of X-ray images of bags not containing a prohibited item. Now, when putting the results in a relationship with the everyday objects task, a very interesting relationship was found in that the screeners who achieved a high performance in the everyday object task were also the ones who generated the fewest false alarms in the simulated X-ray baggage

[1] If an alarm incorrectly indicates the presence of a potentially illicit good while no such good is hidden in the bag, this is considered a false alarm.

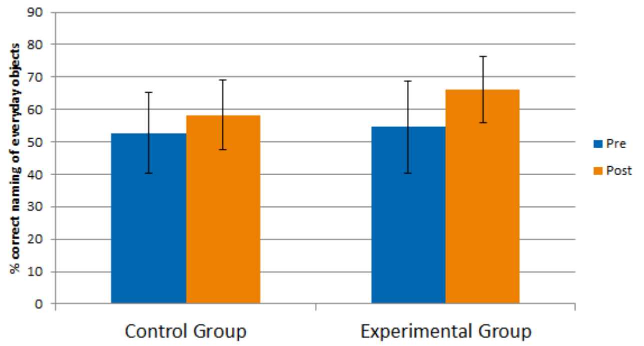

Figure 4: Means for percentage of correct naming of everyday objects for the experimental and control group before (blue) and after (orange) the experimental group conducted the training.

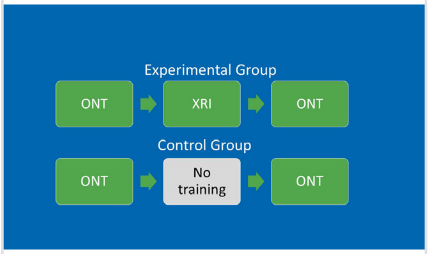

Figure 5: Experimental design used in Experiment 2.

screening task[2]. Therefore, we assume there is a relationship between everyday object knowledge and a low false alarm rate, which is one important component for increasing efficiency.

IS E-LEARNING USEFUL FOR INCREASING EVERYDAY OBJECT KNOWLEDGE?

So far, we have seen that X-ray screeners with good knowledge of everyday objects make fewer false alarms when visually inspecting X-ray images, which in turn should lead to less secondary searches. However, we have also seen that there is some variation in that some screeners already have quite a good knowledge of everyday objects compared to other screeners. Especially X-ray screeners with less experience on the job and therefore less knowledge of everyday objects might highly profit from an everyday object training. Hence, the aim of the second experiment was to test if novices (no work experience in X-ray screening) can recognize harmless everyday objects in X-ray images and whether this can be trained effectively and efficiently using e-learning.

To this end, 30 novices were split into two equal groups (experimental and control group). Both groups were tested twice using an adapted version of the test conducted in Experiment 1, i.e. the X-ray object naming test (X-Ray ONT, without prohibited items). The difference between the groups was that the experimental group conducted an e-learning course on everyday objects (the X-Ray Introduction, XRI, see info box) while the control group did not receive any training between the two ONT tests as shown in Figure 5.

Results showed that in the first test, the groups were equally good at naming everyday objects (Control Group: 53%, Experimental Group: 55% of correctly named everyday objects). This means that novices were able to recognize half of the everyday objects in X-

X-RAY INTRODUCTION (XRI)

The XRI is an e-learning course for beginners teaching the look of harmless everyday objects in cabin baggage and air cargo, which is provided by CASRA (Figure 6). The XRI is recommended for people who have not worked with X-ray images before. In a short introduction, X-ray technology and the meaning of colors in X-ray images are explained. In the following nine sessions, 90 harmless everyday objects are introduced in three steps. First, harmless everyday objects are shown as X-ray images together with a photograph. As objects sometimes look very different in X-ray images compared to reality, this is an important step to learn what everyday objects look like in X-ray images. Second, X-ray images of passenger bags containing these everyday objects are displayed one after the other, giving the trainees the opportunity to view the introduced objects within a realistic context. Then, trainees have to find each learned everyday object in X-ray images of passenger bags by clicking on the object indicated in the instruction. Participants can also monitor their learning progress in form of a self-evaluation exercise.

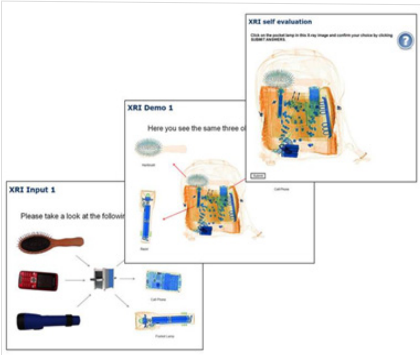

Figure 6: Screenshots of the XRI: Input, demo, and self-evaluation.

[2] There was a linear relationship between % correctly named harmless everyday objects and false alarm rate, i.e. the more everyday objects screeners could name correctly, the lower was their false alarm rate (r(13) = -.471, p=.038) for more information see ([11]).

ray images before having received any training. While this result on the one hand is encouraging, it means on the other hand that many everyday objects cannot be recognized in X-ray images very well without training. Interestingly, the group that received training was able to increase the percentage of correct naming already by about 12% only after 2.5 hours of training as shown in Figure 4.

OPERATIONAL RELEVANCE

From an operational point of view, a low false alarm rate is desirable to increase and maintain the efficiency of the security screening process. In this article, we shed light on the role of knowledge of harmless everyday objects in X-ray screening i.e. whether the knowledge of everyday objects can increase efficiency by decreasing false alarm rates, and whether it can effectively be trained. Experiment 2 showed promising results on e-learning as an effective tool for building knowledge of harmless everyday objects in X-ray images. This is certainly important for initial training so that X-ray screeners can work efficiently as early as possible. Although on average certified screeners reached high values in everyday object naming in Experiment 1, there was still substantial variation between them. Therefore, it could make sense to run additional studies to see whether everyday object recognition training should be implemented not only for initial but also for recurrent training. To sum up, increased knowledge of harmless everyday objects, in combination with sound visual knowledge of prohibited items, could have positive effects on the effectiveness and the efficiency of airport security screening.

REFERENCES

[1] Schwaninger, A., & Hofer, F. (2004). Evaluation of CBT for increasing threat detection performance in X-ray screening. In K. Morgan and M. J. Spector (Eds.), *The Internet Society 2004, Advances in Learning, Commerce and Security* (pp. 147-156). Wessex: WIT Press. [Download PDF]

[2] Hardmeier, D., Hofer, F., & Schwaninger, A. (2006a). 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, 2006*, 393-397. [Download PDF]

[3] Michel, S., de Ruiter, J., Hogervorst, M., Koller, S., Moerland, R., & Schwaninger, A. (2007). 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.* [Download PDF]

[4] Schwaninger, A., Hofer, F., & Wetter, O. (2007). 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.* [Download PDF] .

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

[6] Koller, S., Drury, C., & Schwaninger, A. (2009). Change of search time and non-search time in X-ray baggage screening due to training. *Ergonomics, 52*(6), 644-656.

[7] Halbherr, T., Schwaninger, A., Budgell, G., & Wales, A. (2013). Airport security screener competency: a cross-sectional and longitudinal analysis. International *Journal of Aviation Psychology, 23*(2), 113-129. [Download PDF]

[8] Michel, S., Mendes, M., de Ruiter, J., Koomen, G., & Schwaninger, A. (2014). Increasing X-ray image interpretation competency of cargo security screeners. International *Journal of Industrial Ergonomics*, 44, 551-560.

[9] Sterchi, Y., & Schwaninger, A. (2015). A First Simulation on Optimizing EDS for Cabin Baggage Screening Regarding Throughput. *Proceedings of the 49th IEEE International Carnahan Conference on Security Technology, Taipei Taiwan, September 21-24, 2015,* 55-60. [Download PDF]

[10] Gkritza, K., Niemeier, D., & Mannering, F. (2006). Airport security screening and changing passenger satisfaction: An exploratory assessment. *Journal of Air Transport Management*, *12*(5), 213–219.

[11] Hättenschwiler, N., Michel, S., Ritzmann, S., & Schwaninger, A. (2015). A First Exploratory Study on the Relevance of Everyday Object Knowledge and Training for Increasing Efficiency in Airport Security X-ray Screening. *Proceedings of the 49th IEEE International Carnahan Conference on Security Technology, Taipei Taiwan, September 21-24,* 2015, 25-30. [Download PDF]

Science & Security – CASRA Newsletter
ISSN 1664-5715
Published three times a year

Editors:
Dr. Diana Hardmeier
Prof. Dr. Adrian Schwaninger

Editorial Managers:
Alex Kunz
Flavia Kohler

CASRA
Thurgauerstrasse 39
8050 Zurich, Switzerland

Phone +41 (0)43 336 01 01
Fax +41 (0)43 336 01 00
E-mail info@casra.ch
Web www.casra.ch