SCIENCE & SECURITY November 2013

CASRA NEWSLETTER – ISSUE 5

In its second year the CASRA newsletter continued to experience growing attention and interest as shown by the increasing number of readers. In this issue we address two new topics that seem to be of importance to a wide audience.

The first article explains usability research and testing, which is essential to make software easy to use. It shows how CASRA evaluates its software continuously regarding usability in order to provide best-in-class solutions which are user-friendly.

The second article looks at key performance indicators, a topic of high importance for everyone dealing with the measurement and management of security performance.

We hope this newsletter is again useful and insightful for you and we welcome any feedback you might have regarding topics you would want us to address next year.

CASRA wishes you a nice advent season and a Happy New Year.

Prof. Dr. Adrian Schwaninger

Dr. Diana Hardmeier

TOPICS IN THIS ISSUE:

RESEARCH PUT ACROSS

USABILITY - THE KEY TO GETTING THINGS DONE RIGHT ON THE FIRST ATTEMPT

These days, people constantly use software in all aspects of life - and that includes people who claim not to own a computer. We use software when we purchase train tickets, when we browse the World Wide Web, when we set the alarm clock, and so on. With software so prevalent in our lives, it is very important to make software easy to use for all sorts of people who may interact with it.

SECURITY IN PRACTICE

KEY PERFORMANCE INDICATORS FOR SECURITY SCREENING

You can only manage what you can measure. This article summarizes the results of a survey conducted in different countries on key performance indicators (KPIs) used in airport security. Based on the survey results and CASRA's research experience a set of KPIs for security screening is proposed.

USABILITY - THE KEY TO GETTING THINGS DONE RIGHT ON THE FIRST ATTEMPT

Text: Philipp Sury

These days, people constantly use software in all aspects of life - and that includes people who claim not to own a computer. We use software when we purchase train tickets, when we browse the World Wide Web, when we set the alarm clock, and so on. With software so prevalent in our lives, it is very important to make software easy to use for all sorts of people who may interact with it.

Of course, when attempting to make software easy to use, the big question always is how it can be made easy to use. An entire field of research, called usability research, exists today that answers exactly this question (see [1]). A multitude of professionals coming from diverse fields such as biology, psychology, business administration, and computer science work together to make software easy to use.

The psychologists' approach to usability research is based on a strong understanding of typical human traits and how the human mind works. This knowledge makes psychologists play a key role when it comes to making software fit to human beings rather than attempting to make human beings match to a certain piece of software.

International norms have been designed (ISO, 1999, 2008) to define what makes a good usability in the software. In short, it is these three aspects:

- Effectiveness
- Efficiency
- Satisfaction

The effectiveness of software describes whether human beings are able to reach their goals - e.g. setting the alarm clock correctly, and board the train with a valid ticket. The efficiency of software describes how quickly human beings reach their goals. If someone spends half an hour at an unfamiliar ticket vending machine with the train about to leave in the final minutes, one may experience feelings of stress and anger even if finally fetching a valid ticket and boarding the train before it leaves the station. Last but not least satisfaction is heavily based on effectiveness and efficiency as it can be safely assumed that overly complex software will leave people rather unhappy, but on top of that satisfaction covers all aspects of a user's attitude on a software product in a general way. One can even go so far as to extending the concept of satisfaction to aspects like how the experience was at the store prior to the purchase, the availability and friendliness of customer support, a concise and well written handbook, and so on (see [2]).

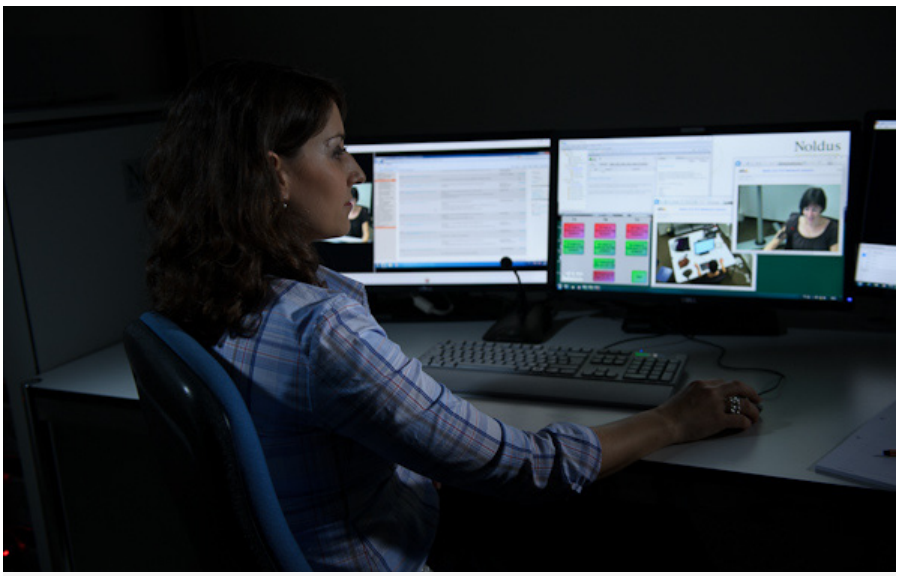

Figure 1: Usability Lab at the University of Applied Sciences and Arts Northwestern Switzerland (Picture: Geri Krischker)

CASRA AIMS HIGH

When it comes to software, the easiness of use has always been very important for CASRA ever since the first version of X-Ray Tutor (XRT) was released more than ten years ago. CASRA employs numerous psychologists, who also received training in the field of usability during their studies, and these psychologists' inputs are being used in daily product design. As CASRA is continuously expanding its range of security software products and transforms them into a software suite, these inputs become even more important to keep software complexity for the end user as low as possible in order to make it user friendly. However, even the most skilled usability professionals automatically get too entangled in a software project and lose their objectiveness, so they require input from outside people. Dumas & Redish (see [3]) call this asking the "How" and "Why" questions that lead to a better understanding of how people unfamiliar with a certain

piece of software interact with it. For that purpose, CASRA uses its connections to the University of Applied Sciences and Arts Northwestern Switzerland for an independent outside view on its software products. CASRA aims at enabling people to use its software suite without detailed study of a user manual, so running usability studies together with the university enables CASRA to progressively advance the quality of its software suite.

FEEDBACK FROM USERS

In order to continue to improve the usability of its software, CASRA lately ran an extensive usability study on the administration panel in its software suite. CASRA wanted to know how easily both experts and novices are able to use the functions in the administration area. For that purpose, the usability study was run with two sets of users:

- A usability questionnaire was sent to 18 customers working as administrators with CASRA's software suite.
- A total of ten college students without any prior work experience with CASRA's software suite performed a usability test with tasks comparable to the daily work of an administrator.

To assess the ease of use of CASRA's software suite as it is experienced by customers, the IsoMetrics usability questionnaire was used (see [4]). The IsoMetrics questionnaire provides a series of statements such as e.g. "I can interrupt any dialog at any time" or "I can tell straight away which functions are invoked by the various menu items" that are meant to assess the relevant aspects of usability as defined by the aforementioned ISO (2008) norm. Customers filling out the questionnaire indicated their agreement or disagreement with each statement on a scale of 1 to 5 with a grade of 1 describing a predominant disagreement and a 5 demonstrating a predominant agreement.

For the usability test (for more information see [5]), college students were invited to a usability lab and asked to perform tasks in the administration area such as adding new users to the training, defining their own set of training course, verify training times, send messages to trainees via the Infoboard application, etc. While the students performed these tasks, they were recorded on film, the content of their computer screen was recorded with a screen capturing device, and the position of their eyes was observed with eye tracking hardware. Using these methods, a multitude of data was collected that allowed the researcher to observe the point of a user's attention (eye tracker), problems encountered while using the software suite (screen capture), and displays of frustration and satisfaction (movie showing a user's face).

After collection, data from the IsoMetrics questionnaire (see [4]) and from the usability test (see [5]) was statistically evaluated to receive an overall picture of the strengths and weaknesses of the newest CASRA software suite.

THE PROFESSIONALS ARE HAPPY

A total number of nine valid questionnaires was sent back from administrators with CASRA's software suite and included in the statistical analysis.

The results gained from the IsoMetrics questionnaire (see [4]) showed that customers reacted very favorably toward the CASRA software suite when it comes to its ease of use (see Fig. 2).

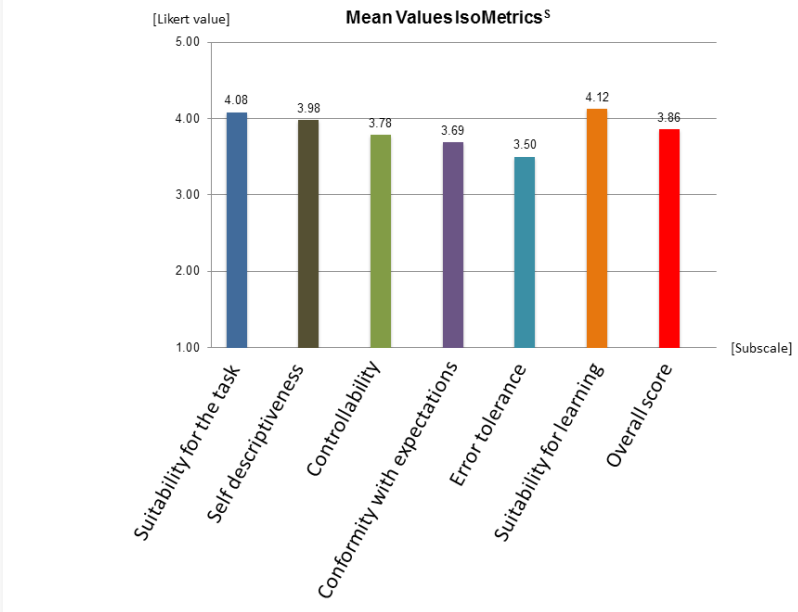

Figure 2: Mean values obtained in nine valid IsoMetrics questionnaires filled out by customers to assess the usability of the administration area in CASRA's software suite. A score of 1 reflects the worst possible result, a score of 5 the best possible result. No values for the scale "suitability for individualization" were collected as CASRA's software suite currently does not offer any options in this regard.

All aspects were rated above a grade of 3 indicating that customers have an overall good impression. Customers gave the highest rating to the scale "suitability of learning", i.e. how easily somebody new to the software suite is able to learn how to use it. However, all participants used the CASRA software suite already prior to filling in the questionnaire, so it is uncertain how accurate their memories are of the time when they first started to use it. Nevertheless, some participants, who are still using an older version of CASRA's software suite, got a free demo access to the latest software suite together with the questionnaire and showed the same favorable reaction. Therefore one can conclude that existing customers find it easy to upgrade from an older to the most current version.

EASY ENHANCEMENTS

For the usability test, college students worked with the CASRA software suite without any prior knowledge on how to use it. They were presented with tasks such as adding new users, define courses, verify training times, messaging other users, etc. The reasoning is that if college students without any prior knowledge of the CASRA software suite can successfully work with it, then professional first-time users should find it even easier. Furthermore, experienced professional users will benefit from the increased comfort experienced after fixing the most common difficulties for inexperienced users.

The first problem encountered by the inexperienced college students was that four out of ten failed to add the new user account to the correct group. They assumed that a new user account is automatically added to the group they currently work with and failed to see the drop down menu in the window where the desired user group can be selected. This difficulty can easily be fixed by setting different default values, i.e. making the current group the default selection in said drop-down box whenever a new user is created. If the default group is the one the user wants (which will be mostly the case), then the user does not have to pay attention to the group selection anymore and can simply move on with the task. If the default group happens to be incorrect, the user can still change the group in the drop-down menu. As the time of this writing, this enhancement has already been implemented in CASRA's software suite.

The second problem was the system feedback when setting a password for a new user account. CASRA takes software security seriously and enforces the use of high quality passwords, but the warning messages in the password setting process were irritating for some users albeit correct. A quick redesign of message content and repositioning will solve this problem.

A third problem was that the college students failed to add users to the system by the means of a Microsoft Excel template file. Post-test interviews indicated that they did not realize where they could download the necessary template, so renaming the download button will be a big help to solve this problem.

Last but not least, the college students took a long time to notice the load button in the reports section although they eventually all succeeded in loading the XRT training data. It was found that said button breaks the conventions for the appearance of buttons inside CASRA's software suite, so optically redesigning said button to match these conventions will be an easy solution to this difficulty.

SO EASY EVERYONE CAN USE IT

Besides the few tasks that were difficult for the inexperienced college students to solve, there were also a number of tasks even the inexperienced college students found very easy.

The first was setting up a user account for two-step verification with an mTan code for mobile users (see Fig. 3). While two steps authentication is not new as a concept, it is getting increasingly important in today's world of mobile and cloud computing to prevent hacker attacks. Especially in the case of people picking bad passwords and falling victim to lateral attacks, two-step verification is an important line of defense. CASRA is therefore glad to observe how quickly and easily the col-

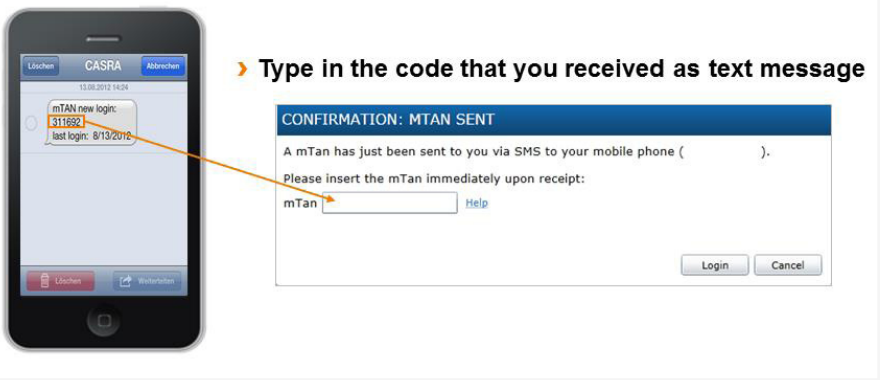

Figure 3: For added security, an mTan code can be sent to a user's phone during the login process.

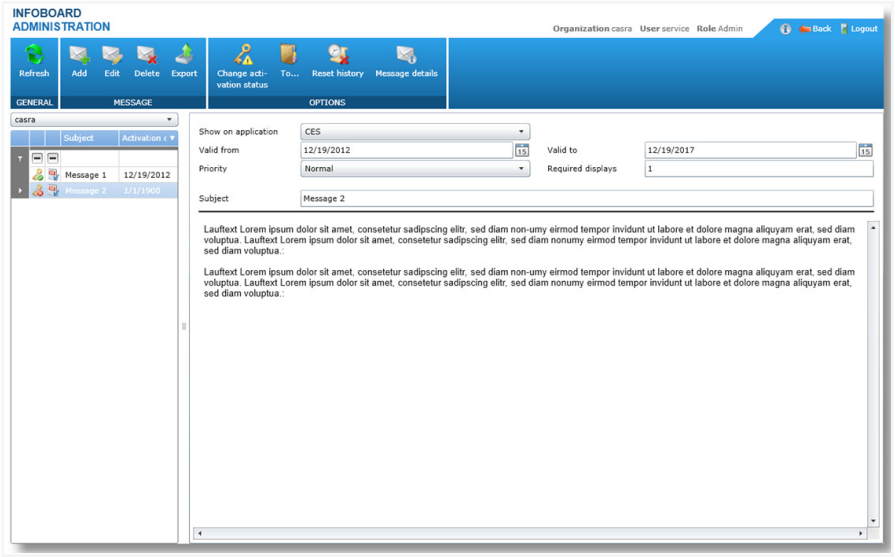

Figure 4: The Infoboard application within the new CASRA software suite

lege students interacted with its two-step authentication process.

The second task that went smoothly was logging on as a screener and read the waiting Infoboard messages. This task was easily solved by all participants and stands as a confirmation for the Infoboard to be very useful for well-managed delivery of information to many employees in a short amount of time (see Fig. 4).

Last but not least, the college students also found it easy to deactivate a user account, be it to temporarily block somebody's access to CASRA's software suite (e.g. when somebody is on an extended leave from the company) or if for whatever reason somebody's access needs to be revoked.

CONTINUOUS STEPS ON AN EXCITING ROAD

CASRA is strongly committed to high usability of its software. Professional usability tests are essential to examine whether users can effectively and efficiently work with a piece of software and how satisfied they are with the interaction process. Usability tests reveal software strengths that can still be honed and weaknesses that need improvement. The above mentioned usability study has shown that CASRA's customers are happy with its current software suite and see it again as an improvement compared to earlier versions. Tests with inexperienced students, however, helped to identify a number of areas for further improvement that require relatively small adaptations, but can greatly increase the ease of use for absolutely novel users. While inexperienced college students are not at all the target audience for CASRA's software suite, insights gained from these tests will also be of benefit for advanced and experienced professionals. CASRA's goal is to build a software suite that requires no prior knowledge, no extensive training, and no tiresome flipping through dozens of pages in a handbook. For that purpose, a scientific and systematic approach to systematic usability testing and enhancement has played an important role in CASRA's past and will continue to do so in the future.

REFERENCES

[1] Rosson, M. B., & Carroll, J. M. (2002). *Usability Engineering*. San Francisco: Morgan Kaufmann Publishers.

[2] Groner, R., Raess, S., & Sury, Ph. (2008). Softwareergonomie und Usability. In B. Batinic (Ed.): *Medienpsychologie*. Heidelberg: Springer.

[3] Dumas, J.S., & Redish, J.C. (1999). *A practical guide to usability testing*. Exeter: intellect.

[4] Gediga, G., Hamborg, K.-C., & Düntsch, I. (1999). The IsoMetrics usability inventory: an operationalization of ISO 9241-10 supporting summative and formative evaluation of software systems. *Behavior & Information Technology, 18*, pp. 151-164.

[5] ISO (1999). Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability (ISO 9241-11:1998); German version EN ISO 9241-11:1998.

[6] ISO (2008). Ergonomics of human-system interaction - Part 110: Dialogue principles (ISO 9241-110:2006); German version EN ISO 9241-110:2006.

KEY PERFORMANCE INDICATORS FOR SECURITY SCREENING

Text: Marcia Mendes, Alex Kunz & Adrian Schwaninger

This article summarizes the results of a survey on key performance indicators (KPIs) related to airport security screening and contains a recommendation of KPIs for the security performance domain.

An airport security screening checkpoint can be regarded as a system in which humans and technology interact based on defined processes producing certain outputs. An important output is the detection of prohibited items to prevent that objects such as guns or explosive devices can be used to threaten air transportation. Other outputs such as throughput or passenger satisfaction are often taken into account as well. KPIs are measurements that are used to gauge the performance of a system and judge whether an expected output is successfully generated or not.

MISSING STANDARDS FOR KPIs

Unfortunately, there are no international standards for such KPIs which make it difficult to optimize system performance and compare the effectiveness of different security checkpoint solutions. To shed light on the use and implementation of KPIs, Hofer, Mendes and Schwaninger (see [1]) have conducted a survey in five different countries. Information about KPIs was provided on an anonymous basis from private security companies, representatives of appropriate authorities, and airport operators. Each organisation listed between six and eleven KPIs. The data was categorized as shown in Figure 1 into performance domains and working area domains. The results of this survey can be summarized as follows:

- **Large differences exist regarding which KPIs are used and how they are or should be measured.** There is considerable variance in the frequency of measurement and aggregation level of the KPIs between the different organisations. Some KPIs are measured quarterly, some yearly, others monthly. In addition, data aggregation level also differs for the same performance indicators in a large amount, e.g. mean value per month, quarter, trimester, semester or year.
- **Most indicators refer to cabin baggage and passenger screening.** This is not surprising since this working area domain has the most external visibility. However, other working area domains like cargo or staff/crew screening are of similar importance to ensure aviation security.
- **The majority of indices belong to the security performance domain.** Performance domains such as facilitation or customer service, which are relevant from an economic point of view, are less frequently measured.

RECOMMENDATIONS FOR KPIs IN THE SECURITY PERFORMANCE DOMAIN

The conducted survey has shown that it would be helpful and of great value to set international standards for the use of KPIs. Based on the findings of the conducted survey and experience of CASRA in many research projects at a large number of airports in different countries, a set of KPIs can be recommended for the security performance domain as described in the following sections.

COVERT TEST RESULTS

To examine the effectiveness of security screening, covert tests should be conducted by trained personnel who carry prohibited items (e.g. knives or mock improvised explosive devices (IEDs)) on the body or in passenger baggage. These tests should be conducted by or on behalf of the appropriate authority for external quality control and also by airports and/or security companies for internal quality control and train-

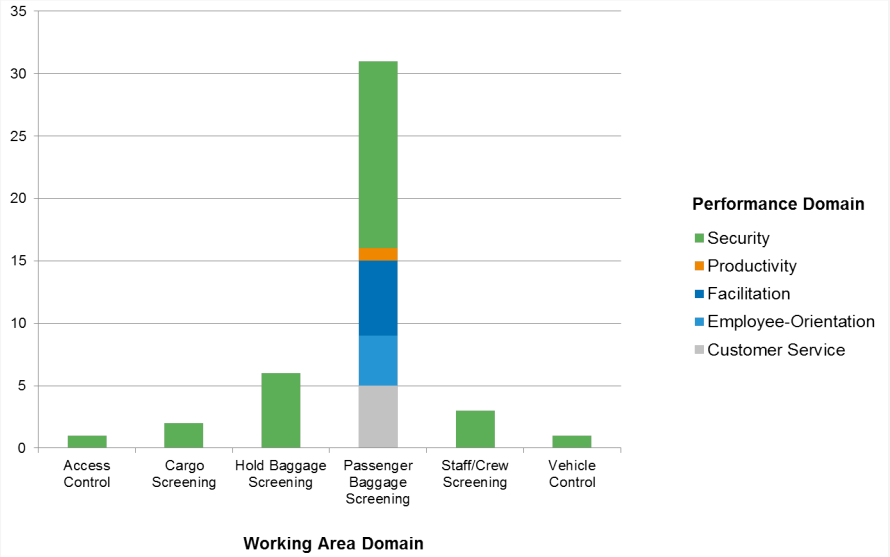

Figure 1: Number and types of KPIs

ing purposes. Covert tests conducted at screening checkpoints can be used to verify whether screeners conduct the security control procedures correctly, whether screeners are able to detect prohibited items while working at a security checkpoint, and whether screeners react appropriately once a prohibited item has been detected.

Covert test results depend on several factors, including:

- Screeners' competencies (i.e. knowledge, skills, and/or abilities)
- Screener and team motivation, attention and awareness
- Distraction at the moment of testing
- How and how often covert tests are conducted
- External factors (e.g. staffing, test difficulty, technology and checkpoint design)

Conducting covert tests on a regular basis is very important in order to measure operational security performance and increase motivation, attention, and awareness of screeners. So far, only a few studies have addressed covert tests scientifically (e.g. [2] & [3]). Further research is needed in order to provide scientifically based guidance on how and how often covert tests should be conducted.

OVERT TEST RESULTS

Overt tests should be conducted by trained personnel for internal quality control by x-raying realistically prepared passenger baggage in the presence of security screeners and asking them to describe the content of the X-ray image. Overt tests can be used to verify whether screeners are able to recognize prohibited and everyday objects in X-ray images, and describe them appropriately. Overt tests should also be conducted by appropriate authorities for external quality control.

Overt test results depend on several factors, including:

- Screeners' competencies (i.e. knowledge, skills, and/or abilities)
- Nervousness at the moment of testing

Distraction at the moment of testing and insufficient motivation or awareness can almost be ruled out because tests are overt. Conducting overt tests on a regular basis in addition to covert tests is important in order to better determine the reason why screeners fail at covert tests, to increase motivation and awareness and to reduce effects of nervousness. As for covert tests, further research is needed in order to provide scientifically based guidance material on how, how often and by whom overt tests should be conducted.

IMPLEMENTATION OF ON-THE-JOB TRAINING

On-the-job training (OJT) provides immediate feedback to the person undergoing the training, which is seen as a key factor to improve detection performance. OJT should be conducted to ensure that screeners achieve the required performance and are able to translate their acquired theoretical knowledge into practice. This type of training should be conducted by screening providers when overt and/or covert tests have been failed and to prepare screeners for overt and covert testing as part of their initial and recurrent training program.

RECURRENT X-RAY IMAGE RECOGNITION TRAINING PERFORMANCE

Recurrent training and testing of X-ray image recognition is mandated by regulations of the European Commission (see [4]). Appropriate authorities, airports and/or screening providers should monitor whether each screener conducts the required training hours. Exposing trainees to six hours of image recognition training on one day is neither useful, nor effective. Instead, it is recommended to

split up the six training hours required every six months into multiple sessions. Although the indicator of six hours every six months is relatively easy to monitor, it does not provide information on the quality of the image recognition training itself. It is therefore recommended to define additional indicators for X-ray image recognition training and testing, to monitor performance improvement over time, detect weaknesses and provide purposive training to address the areas in need of improvement. For more information on important aspects to consider regarding computer-based training see [5].

Computer-based training systems should have administration functions where training managers can easily retrieve training and test results. Moreover, it is recommended to use a system featuring individually adaptive training algorithms. If this is not available, the computer-based training should at least allow managers to adapt the settings of training- and testing sessions manually, allowing differentiation and customization of training and testing, taking into account individual training needs, adjusted to the screener level performance of the X-ray operator.

CERTIFICATION RESULTS

The European Commission Regulation 185/2010 specifies that persons performing certain screening tasks have to be subject to an initial certification or approval process and be re-certified periodically. X-ray image interpretation exams and theoretical exams can be conducted by using computer-based tests. When implemented properly, this allows measuring screener competencies (i.e. knowledge, skills, and/or abilities) needed for good operational performance in a reliable, valid and standardized way. Detailed guidance material on X-ray image interpretation exams can be found in Schwaninger et al. (see [6]). Because computer-based tests cannot directly measure the actual behaviour of screeners at a security checkpoint (which is also dependent on motivation, awareness, attention and other factors), covert and overt tests as well as on-the-job training should be conducted in combination.

THREAT IMAGE PROJECTION (TIP) PERFORMANCE

Threat image projection (TIP) is a software function of X-ray machines, which allows to project fictional threat items (FTIs) into real X-ray images of passenger bags during routine X-ray screening operations. While TIP is not considered useful for training, it can be valuable for increasing motivation, attention and awareness if implemented properly combined with covert and overt testing procedures (see [7]). By means of a TIP-data analysis-software, hit rates, false alarm rates and throughput data my be analyzed. For meaningful analyses, individual performance data should be aggregated over several months and reliability should be analyzed statistically (for more information see [8]).

FURTHER CONSIDERATIONS

Information on KPIs should be stored and shared between entities conducting external and internal quality control wherever possible, in order to compare and validate results and gain information on how to improve security screening performance. Moreover, it is beneficial when airport operators and security screening providers work in collaboration when developing KPIs (e.g. [9]). Finally, it should be noted that for work motivation it is much more effective to use KPIs for bonuses instead of maluses if KPIs are related to financial consequences.

For more information regarding the implementation of key performance indicators and measurement methods please contact CASRA [info@casra.ch].

REFERENCES

[1] Hofer, F., Mendes, M., & Schwaninger, A. (2011). Towards a common understanding of key performance indicators in aviation security. Version 2100-10-05.

[2] Schwaninger, A. (2009). Why do airport security screeners sometimes fail in covert tests? *Proceedings of the 43rd IEEE International Carnahan Conference on Security Technology, Zurich Switzerland, October 5-8, 2009.*

[3] Wetter, O., Hardmeier, D., & Hofer, F. (2008). Covert testing at airports: Exploring methodology and results. *Proceedings of the 42nd Carnahan Conference on Security Technology, Prague, October 13-16, 2008.*

[4] European Comission. (2010). Commission regulation (EU) No 185/2010 of March 2010 laying down detailed measures for the implementation of the common basic standards on aviation security. *Official Journal of the European Union, 53,* 1–55.

[5] Schwaninger (2011). Computer-based training: advantages and considerations. *Aviation Security International, 17*(6), 18-23.

[6] Schwaninger, A., Bridges, A., Drury, C., Durinckx, F., Durrant, P., Hodge, T., Hofer, F., Jongejan, R., Maguire, R., McClumpha, A., Neiderman, E., Steinmann, C., & Wüest, W. (2006). Principles and requirements for assessing X-ray image interpretation competency of aviation security screeners. *White Paper, International Aviation Security Human Factors Technical Advisory Group (InterTAG), Competency Assessment Working Group (CAWG).*

[7] Schwaninger, A. (2006b). Airport security human factors: From the weakest to the strongest link in airport security screening. *Proceedings of the 4th International Aviation Security Technology Symposium, Washington, D.C., USA, November 27 – December 1, 2006,* 265-270.

[8] Hofer, F., & Schwaninger, A. (2005). Using threat image projection data for assessing individual screener performance. *WIT Transactions on the Built Environment, 82,* 417-426

[9] Wetter, O. E., Fuhrmann, H., Lipphardt, M., & Hofer, F. (2011). Bringing adversaries together: The importance of a common management-level approach in complex work domains. In L. D. Sanson (Ed.), *Proceedings of the 45th International Carnahan Conference on Security Technology. Mataró, Spain: Institute of Electrical and Electronics Engineers.*

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

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