Comparing e-learning and blended learning for threat detection in airport security X-ray screening

Stefan Michel, Nicole Hättenschwiler, Melina Zeballos, Adrian Schwaninger
School of Applied Psychology
University of Applied Sciences and Arts Northwestern Switzerland (FHNW)
Olten, Switzerland
and
Center for Adaptive Security Research and Applications (CASRA)
Zurich, Switzerland
Email: stefan.michel@fhnw.ch

Abstract—This study investigated the effectiveness of e-learning and blended learning for airport security officers (screeners) aimed at increasing their detection of novel prohibited items in X-ray images of passenger bags. 80 screeners from a large European airport conducted a simulated X-ray baggage screening task (SXBST). Based on hit- and false alarm rates, d-prime, age, gender and job experience, four equivalent screener groups were created. The first group (EL1) conducted six different e-learning modules. The second group (EL2) repeated the e-learning modules once with a predefined time interval of one week in-between. The third group (BL) participated in an asynchrony blended learning course where all participants first conducted the six e-learning modules followed by classroom training. The control group (CG) did neither conduct e-learning nor participate in the blended learning course. All four groups conducted the SXBST again after these interventions. Statistical analyses of the hit rates for novel prohibited items showed a significant main effect of group (EL1, EL2, BL, CG). Post-hoc tests showed that the increase of the hit rate for EL1 vs. CG, EL2 vs. CG and BL vs. CG was significant. Therefore, all interventions helped to increase the detection of novel prohibited items in X-ray images of passenger bags. Additionally, the hit rate for BL was significantly higher compared to EL1, which underlies the benefit of blended learning as a combination of classroom and online-training. The repetition of the e-learning modules (EL2) led to a marginally significant increase of the hit rate compared to EL1. The difference between EL2 and BL was not significant. This suggests that when training duration is controlled, the advantage of blended learning compared to e-learning vanishes. In addition, training did not impact false alarm rate as the main effect of group was not significant. Analysis of response times showed no speed-accuracy tradeoff but screeners needed significantly more time to analyze target-absent images than target-present images and this effect was independent of group. Overall, these findings suggest that e-learning as well as blended learning are effective to improve screeners’ detection of novel prohibited items in X-ray images of passenger bags. Implications for X-ray image interpretation training for screeners are discussed.

Keywords—airport security, e-learning, blended learning, threat assessment

I. X-RAY IMAGE INTERPRETATION TRAINING IN AIRPORT SECURITY

One of the main tasks of airport security officers (screeners) to prevent terrorist attacks is to detect prohibited items in X-ray images of passenger bags. To fulfill this task, screeners have to learn which items are prohibited and what they look like in X-ray images. Different methods to train screeners are in use at airports worldwide. The goal of this study was to investigate the usefulness of two methods: e-learning and blended learning.

Computer-based training for X-ray image interpretation (CBT) was investigated in different studies (e.g. [1], [2], [3]). One widely used CBT for X-ray image interpretation is X-Ray Tutor (XRT, for further information see [4], [3]). Within one learning session of XRT, screeners are exposed to hundreds of different X-ray images of passenger bags and many prohibited items. During training, screeners learn how prohibited items look like in X-ray images of passenger bags in different rotations, different degrees of superposition as well as in bags with various complexities. The computer provides feedback about the correctness of the answer, which is an important aspect for the learning success ([4], [5]). Studies could show that training with XRT led to an increase of detection performance as well as a faster detection of threat items in passenger bags (e.g. [6]). Typically, analyzing target absent images needs more time (e.g. [7], [8]). Fast response times, low false alarm rates and other factors (e.g. the availability of an explosive detection systems, EDS) are relevant for a good passenger throughput [9].

One important aspect of a CBT for X-ray image interpretation training (e.g. XRT) is to keep the image library of prohibited items up to date regarding new and emerging threats. To this end, a systematic threat assessment (STA) is important where different sources (surface web, deep web, social media, radical propaganda magazines etc.) are systematically searched for potential threat scenarios. Subsequently, these threat scenarios are evaluated and assessed with regard to feasibility, damage potential and possible mitigation measures [10]. Potential prohibited items are rebuilt in cooperation with police organizations, systematically recorded with X-ray machines and
integrated into CBT. Especially prohibited items which can be detected by visual properties (e.g. shape, color, structure) are eligible to be used in such a CBT for X-ray image interpretation training. Whenever additional information about prohibited items and their potential unlawful usage is relevant for the detection (e.g. how improvised explosive devices (IEDs) can be hidden in electronic devices or in toys) other methods of learning could be considered like for example e-learning or blended learning.

E-LEARNING

Different studies investigated the usefulness of e-learning in different working fields such as health care (e.g. [11]) and higher education (e.g. [12]).

E-learning can be defined as "the use of computer network technology, primarily over an intranet or through the internet, to deliver information and instruction to individuals" [13]. Beside e-learning different terms for the above-mentioned definition are used in the literature: computer-based learning, on-line learning, distributed learning, or web-based training [13]. In addition, reference [14] used the term Modular Distance Learning (MDL) which refers to "packaged learning modules designed to maintain competencies and to excel at relevant knowledge and skills…". The different terms and broader definitions are interesting but e-learning is based on reference [13] the term of choice in organizations. E-learning is also the term used in this article for learning content provided in modules, where pictures, X-ray images of prohibited items and contextual information are combined.

The advantages of e-learning are discussed in several scientific articles and can be summarized as follows: With e-learning the content is accessible to all stakeholders and a third of e-learners do most of their e-learning in the comforts of their homes [14]. In general, the benefits reported for e-learning are satisfaction, cost-effectiveness, accessibility and flexibility (e.g. [15], [16], [17]) as well as that students can own control over the content, place and time of learning, which is reported in a meta-analysis by [18]. Furthermore, reference [18] stated that e-learning can help students gain knowledge and skills faster than traditional instructor-led methods. A disadvantage is mainly the lack of social interaction [19]. For example, reference [20] reported that "some socially constructed barriers to learning may disadvantage learners in online environments". Reference [21] further argued that learning requirements and preferences of each learner tend to be different.

BLENDED LEARNING

Another learning methodology, which is often discussed and also investigated in scientific studies, is blended learning. Blended learning is often associated with simply linking traditional classroom training to e-learning activities [21]. Different methods of how blended learning can be implemented are discussed in the literature. One possibility is asynchronous work where learners conduct the e-learning at their own time and pace outside the classroom and participate then in classroom training later on [21]. In our study we understood the term blended learning as a combination of online and offline learning methods where e-learning is provided over the internet and offline learning happened in a classroom setting. For our experiment we used an asynchronous form where at first e-learning was conducted online and self-paced and screeners participated in the classroom training later on.

Blended learning was investigated in different scientific studies and learning settings (e.g. [22]). The advantages of blended learning range from "more time for tutor support", "improve the learning experience" and "indicates that blended learning not only offers more choices but also is more effective" [20]. References [23] stated that "the current focus on learning technology in traditional universities indicates the significance of blended learning research for both society and economy". Reference [24] noted that blended learning design and implementation are heavily context-dependent, and generalization can be challenging.

The goal of the current study was to investigate the usefulness of e-learning and blended learning in the context of airport security X-ray screening. Specifically, it was investigated whether e-learning and blended learning are useful methods to train screeners detection of novel prohibited items in X-ray images of passenger bags. We hypothesized that e-learning and blended learning are both useful methods to increase screeners detection of novel prohibited items in X-ray images of passenger bags. In addition, we hypothesized that blended learning should be more effective than e-learning because of the positive effects discussed in other domains (e.g. [20]).

II. METHODS

PARTICIPANTS

In total 80 screeners from a large European airport were recruited for this study. After the baseline test, five participants dropped out due to illness and other reasons. Therefore, 75 screeners completed the whole study. A description of the sociodemographic characteristics per group is listed in Table 1.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of participants</th>
<th>Age (M/SD)</th>
<th>Gender (%male/ %female)</th>
<th>Job experience (M/SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control group (CG)</td>
<td>20</td>
<td>49.70/9.13</td>
<td>45/55</td>
<td>7.75/5.43</td>
</tr>
<tr>
<td>Blended Learning (BL)</td>
<td>18</td>
<td>44.55/10.05</td>
<td>56/44</td>
<td>9.39/6.95</td>
</tr>
<tr>
<td>E-Learning 1 (EL1)</td>
<td>20</td>
<td>44.85/8.50</td>
<td>50/50</td>
<td>8.05/4.58</td>
</tr>
<tr>
<td>E-Learning 2 (EL2)</td>
<td>17</td>
<td>45.82/11.66</td>
<td>47/53</td>
<td>9.24/5.67</td>
</tr>
</tbody>
</table>

MATERIALS

A simulated X-ray baggage screening task (SXBST, see Fig. 1) was used to measure detection performances as well as reaction times.

TABLE 1: NUMBER, AGE, GENDER AND EXPERIENCE OF PARTICIPANTS PER SCREENER GROUP
The XBST consisted of 512 X-ray images of passenger cabin bags. 128 different prohibited items from four different categories (guns, knives, IEDs, others) were used. 64 prohibited items were known to the screeners from the CBT for X-ray image interpretation they conduct on a regular basis (XRT). 64 prohibited items were novel, i.e. not part of the CBT used for recurrent training. Those novel prohibited items were identified during a systematic threat assessment (STA) of different sources (e.g. surface web, deep web, social media, radical propaganda magazines etc.), rebuilt and recorded as X-ray images. All prohibited items were virtually merged by experts (former screeners) into X-ray images of passenger bags using validated algorithms [25], [26].

E-Learning modules were developed based on the ECLASS model by [27]. It consists of the following elements: E = Explain; C = Clarify; L = Look; A = Act; S = Share; S = Self Evaluate/Submit.

The e-learning modules contained the following prohibited items: Hidden prohibited items in toys; irritant sprays hidden in everyday objects; modified e-cigarettes; non-metal IEDs; IEDs hidden in electronic devices and hidden knives.

Each e-learning module had the following structure: Introduction and explanation of the threat scenario; explanation of feasibility of the threat scenario; explanation of the damage potential of the threat scenario and importance of the threat scenario for the work routine.

Reference [28] proposed a four stages model for evaluation (1. reaction, 2. learning, 3. behavior, 4. results) which can be adapted to the e-learning environment. Each of the above-mentioned e-learning modules was evaluated on level 2 and 3 with the adapted version of the Training Evaluation Inventory (TEI; [29]) on different aspects: likeliness, difficulty, increased knowledge, attitude and training design.

**PROCEDURE**

First, all screeners conducted the SXBST as a baseline test in a dimly lit room where four screeners could be tested in parallel. During the test, X-ray images were moving on the screen from left to right whereas a blue bar marked the X-ray image to be responded on. Several image-enhancement functions (IEFs; e.g. super enhancement, organic only etc.) and zoom functions were available on an onscreen keyboard. For more information about IEFs see [30]. Additionally, further functionalities (e.g. preview window, belt controller etc.) were available on the screen.

Based on hit- and false alarm rates, detection performance (d'), age, gender and job experience, four equivalent screener groups were created. As interventions, six different e-learning modules and a blended learning course were applied dependent on the experimental group screeners were allocated. Group 1 (EL1) conducted the six above mentioned e-learning modules with a predefined time interval of three days in-between the modules. Group 2 (EL2) first conducted all six e-learning modules with the same predefined time interval of three days in-between and then repeated all e-learning modules once. The blended learning group 3 (BL) conducted the six above mentioned e-learning modules and participated then in a classroom training. Blended learning was used in an asynchronous form where first the e-learning modules were conducted online and self-paced with the predefined time lag of three days between the different modules. On average, one week after completing the last e-learning module, screeners of the blended learning group participated in a classroom training. During the classroom training, first the content of the e-learning modules and open questions were discussed with two airport security experts. Subsequently, prohibited items which were learned within the e-learning modules could be viewed as real (rebuilt) objects. In a second part of the classroom training, half of the screeners concealed these prohibited items into passenger bags and X-rayed them whereas the other half of the participants from the BL group had to search for these items on X-ray images. The subgroups then changed and the same procedure was conducted again. Finally, the learners' experiences and conclusions were discussed again with the airport security experts. At the end, screeners evaluated the blended learning course on the same aspects like the e-learning modules.

The reason for using the EL2 group was the overall training duration, which was on average the same as for the BL group. The control group (CG) did neither conduct the e-learning modules nor participated in the blended learning course. Subsequently, on average one week after the interventions, all four groups conducted the SXBST again. The test was the same as used for the baseline measurement with the exception that screeners had additional answer buttons available (manual search, explosive trace detection [ETD], liquid explosive detection [LED] and alarm). The study design is shown in Fig. 2.
III. RESULTS

PERFORMANCE

Screeners’ responses on X-ray images of passenger bags from the second SXBST were analyzed (see study design in Fig. 2). When screener pressed the Not OK button or one of the special control buttons (manual search, explosive trace detection (ETD), liquid explosive detection (LED) or alarm) for bags containing a prohibited item, the answers were counted as hits. When a bag was harmless and screeners pressed the Not OK button or one of the special control buttons, the answers were counted as false alarms. Only performance results for novel prohibited items are reported here. A one-way ANOVA using hit rate as dependent variable revealed a main effect of group (EL1, EL2, BL, CG) with $F(3,71) = 7.359, p < 0.001$. One-tailed post-hoc tests showed significantly increased hit rates of the three experimental groups (EL1, EL2, BL) in comparison to the control group (CG): EL1 vs. CG ($M_{diff} = 0.07$), $t(38) = 2.149, p = 0.02$; EL2 vs. CG ($M_{diff} = 0.12$), $t(35) = 3.794, p < 0.001$; BL vs. CG ($M_{diff} = 0.14$), $t(36) = 4.596, p < 0.001$. Furthermore, comparing the hit rate for BL vs. EL1 shows that the hit rate for BL is significantly higher than for EL1 ($M_{diff} = 0.07$), $t(36) = -2.167, p = 0.02$. Repeating the e-learning modules once (EL2) led to a marginally significant increase of the hit rate compared to EL1 ($M_{diff} = 0.05$), $t(35) = -1.427, p = 0.08$. The difference of EL2 vs. BL was not significant ($M_{diff} = 0.03$), $t(33) = -0.738, p = 0.23$. See Fig. 3 for more information. In addition, training did not impact false alarm rate as the main effect of group was not significant, $F(3,71) = 1.179, p = 0.32$.

REACTION TIME

A mixed ANOVA with trial type (target absent vs. target present) as within-subjects factor and group as between-subjects factor on response times showed a significant main effect of trial type, $F(1,71) = 138.07, p < 0.001$. There was no main effect of group, $F(3,71) = 0.232, p = 0.87$. There was no significant interaction, $F(3,71) = 0.949, p = 0.42$ (see Fig. 4 for more information). Separate ANOVAs for target present and target absent trials showed that there was neither a main effect of group for target absent, $F(3,71) = 0.173, p = 0.91$. These results speak against a speed-accuracy trade-off.

EVALUATION OF E-LEARNING MODULES

Screeners rated each of the e-learning modules regarding the aspects likeliness, difficulty, increased knowledge, attitude and training design on a 5-point Likert scale where 1 was the minimum and 5 the maximum. For this an adapted version of the Training Evaluation Inventory (TEI; [29]) was used. As can be seen in Fig. 5, all aspects were rated well and no average value was < 4.

RATINGS OF BLENDED LEARNING COURSE

The subjective ratings of the blended learning course are shown in Fig. 6. Screeners rated the same aspects like they did for the e-learning modules. As can be seen in Fig. 6 all ratings were higher than 4.5.
IV. SUMMARY, DISCUSSION AND CONCLUSION

The results of this study showed that e-learning as well as blended learning increased screeners’ detection of novel prohibited items in X-ray images of passenger bags. Because the comparison between BL and EL1 was significant, it could be concluded that blended learning was more effective than e-learning. But it has to be considered that when the training duration was the same on average for EL2 and BL group, no significant difference of the hit rate between those groups was observed. Furthermore, when repeating the e-learning modules once (EL2), a marginally significant increase of the hit rate was observed. Furthermore, when repeating the e-learning modules once (EL2), a marginally significant increase of the hit rate was observed in comparison with the group who conducted the e-learning modules only once (EL1).

In this study we investigated for the first time the usefulness of e-learning and blended learning for screeners to learn novel prohibited items for airport security screening. Similar to the outcome of this study, positive effects of e-learning were reported in the literature in different domains like for example in healthcare and higher education (e.g. [11], [12]). Earlier studies in other domains might suggest that blended learning could be more effective than e-learning [20]. When we controlled the training duration between e-learning and blended learning, this benefit disappeared. Moreover, screeners rated the e-learning modules and the blended learning course on five different aspects. All average ratings for the e-learning modules were > 4 and all average ratings for blended learning were > 4.5 on a 5-point Likert scale where 5 was the maximum. Screeners seemed to like the e-learning modules as well as the blended learning: the difficulty of both learning methods was appropriate; screeners perceived that their knowledge was increased by conducting the e-learning modules as well as the blended learning course; the attitude to both methods was positive and the design of the e-learning modules as well as the classroom training was attractive. Together with the outcome of the detection performance increase, these subjective ratings support the appropriateness of the interventions and the utilized pedagogical concept for the e-learning modules, which was based on the ECLASS model [27].

False alarm rates are important in aviation security because the passenger throughput is depending on it [9]. If screeners are generating many false alarms, bags have to be manually inspected which takes time and could cause longer waiting queues at the security checkpoint. In our study, the false alarm rate was not significantly different between the tested groups. We found no indication that the different interventions (e-learning and blended learning) would have a negative impact on efficiency at the security checkpoint.

Finally, reaction times for target present and target absent images were significantly different. This outcome is consistent with earlier findings (e.g. [7], [8]). In addition, our results speak against a speed-accuracy tradeoff (SAT).

Further studies should take into account different pedagogical concepts of developing e-learning modules as well as blended learning courses to measure which of the concept is the most effective and efficient one to train screeners on the detection of novel prohibited items. Furthermore, more research with larger sample sizes is needed. In addition, covert test data and data from Threat Image Projection (TIP; see [31] for more information) could be considered to evaluate the usefulness of e-learning and blended learning for the detection of prohibited items at the security checkpoint (see for example [32]). To conclude, the tested interventions (e-learning and blended learning) helped to increase screeners’ detection of novel prohibited items in X-ray images of passenger bags. E-learning is especially for larger security organization an effective and also efficient method to train screeners on novel threat items. The efficiency of e-learning can be explained by the fact that e-learning can be conducted self-paced and it is fully scalable and can therefore be provided to a large number of screeners. Whereas, blended learning is much more cost and time intensive because classroom training can only be conducted in relatively small screener groups where airport security experts as teachers are needed. Based on our results, a repetition of the learned content is recommended.

REFERENCES
