International Journal of Industrial Ergonomics 102 (2024) 103598

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

journal homepage: www.elsevier.com/locate/ergon

Performance of X-ray baggage screeners in different work environments: Comparing remote and local cabin baggage screening

Marius Latscha[a,*], Adrian Schwaninger[a], Jürgen Sauer[b], Yanik Sterchi[a]

[a] University of Applied Sciences and Arts Northwestern Switzerland, School of Applied Psychology, Institute Humans in Complex Systems, Riggenbachstrasse 16, CH-4600, Olten, Switzerland

[b] University of Fribourg, Department of Psychology, Rue P.A. de Faucigny 2, CH-1700, Fribourg, Switzerland

ARTICLE INFO

Keywords:
Work environment
Noise
Social stressors
Visual search
Time on task
X-ray baggage screening

ABSTRACT

X-ray screening of passenger baggage at airports is performed in different work settings. In local cabin baggage screening (LCBS), airport security officers (screeners) analyse X-ray images of passenger bags at the checkpoints where they are exposed to noise and social stress from passengers. In remote cabin baggage screening (RCBS), screeners work in remote, quiet, and office-like environments. The primary aim of this study was to compare the screening performance in LCBS and RCBS. In addition, we examined the effects of time on task and task load in both work settings. Using linear mixed models, we analysed threat image projection (TIP) data from 1706 screeners collected over two years (669,168 TIP events). The results showed better detection of prohibited articles (higher hit rate) and longer processing times in RCBS than in LCBS. In both settings, we found a decrease in the hit rate with increasing time on task, and this decrease was stronger when the task load was high. Interestingly, the effects of work setting, time on task, and task load were relatively small compared with the inter-individual differences in performance across screeners.

1. Introduction

Before entering an aircraft, passengers and their baggage are screened for prohibited items (International Air Transport Association, 2019). Airport security officers (screeners) visually inspect X-ray images of passenger baggage in order to decide whether they are harmless or suspicious and need to be analysed in a secondary search using manual bag search, explosive trace detection, and other methods (Dorton and Liu, 2016; Sterchi and Schwaninger, 2015). In this job, it is important to make the correct decisions and to make them promptly. Ensuring no prohibited items are on board the airplane is central to passenger security and safety. However, each false alarm (i.e., sending a harmless bag to a secondary search) negatively impacts passenger throughput (Dorton and Liu, 2016; Sterchi and Schwaninger, 2015). Hence, providing a work environment that ensures high detection of prohibited items and good throughput is highly relevant for the airport security industry. Airports use different work settings to screen baggage. In local cabin baggage screening (LCBS), screeners analyse X-ray images of passenger bags at checkpoints where they are exposed to noise and social stress from passengers (Michel et al., 2014). In remote cabin baggage screening (RCBS), screeners work in remote, quiet, office-like environments (Buser and Merks, 2020; Kuhn, 2017). Although both work settings are applied in airports, little is known about their impact on screening performance. To close this research gap, the primary aim of this study was to compare the screening performance of LCBS and RCBS. In addition, we examined the effects of time on task and task load in both work settings. Using linear mixed models, we analysed large volumes of field data (screener decisions on 642,035 X-ray images of passenger bags) from an international European airport that used both work settings. In the remainder of the introduction, we discuss cabin baggage screening. We then explain LCBS and RCBS. This is followed by a discussion of previous studies on noise and social stress. Finally, we conclude by introducing the present study and its expected results.

1.1. Baggage screening

Baggage screening is a visual inspection task that involves searching and decision making (Koller et al., 2009; Wolfe and Van Wert, 2010). This task is challenging because of low target prevalence, having to search for an unknown target set, the possible presence of multiple

* Corresponding author.
E-mail addresses: marius.latscha@fhnw.ch (M. Latscha), adrian.schwaninger@fhnw.ch (A. Schwaninger), juergen.sauer@unifr.ch (J. Sauer), yanik.sterchi@fhnw.ch (Y. Sterchi).

https://doi.org/10.1016/j.ergon.2024.103598
Received 31 October 2023; Received in revised form 8 May 2024; Accepted 21 May 2024
Available online 11 June 2024

targets, and variation in target visibility (for reviews, see Biggs and Mitroff, 2015; Donnelly et al., 2019). Factors influencing screening performance include image-based factors (Schwaninger et al., 2005; Bolfing et al., 2008), screeners' aptitude and cognitive abilities (Hardmeier et al., 2005; Hättenschwiler et al., 2019; Peltier and Becker, 2017, 2020), and the quality and amount of computer-based training (Halbherr et al., 2013; Koller et al., 2008; McCarley et al., 2004; Schuster et al., 2013). Findings from other studies suggest that screening performance varies with time on task (Basner et al., 2008; Buser et al., 2020, 2023; Ghylin et al., 2007; Meuter and Lacherez, 2016; Rubinstein, 2020) and task load (Buser et al., 2023; Meuter and Lacherez, 2016; Rieger et al., 2021). Specifically, the hit rate decreased with time on task, and this decrease was accelerated when task load was high (Buser et al., 2023; Meuter and Lacherez, 2016). Environmental stressors such as noise (Yu et al., 2015) and social stressors (Marin et al., 2007; Michel et al., 2014) have also proven to be relevant to performance in X-ray baggage screening.

1.2. Centralised image processing, LCBS, and RCBS

While screening cabin baggage for prohibited articles, screeners are typically positioned at an X-ray machine at the checkpoint and receive image input exclusively from this X-ray machine. However, some airports network their X-ray machines and process their images centrally. With centralised image processing (CIP) (Buser and Merks, 2020; Kuhn, 2017), screeners can inspect images from X-ray machines located elsewhere (e.g., in other lanes). CIP allows airports to allocate personnel resources more flexibly and efficiently by eliminating the one-to-one relationship between the screener and the X-ray machine. Moreover, this approach allows for more options regarding where screening takes place: pooled X-ray images can be distributed to analyst stations located in the lanes next to the X-ray machines (LCBS), or they can be analysed remotely in a separate room (RCBS; Buser and Merks, 2020; Kuhn, 2017). Although RCBS is a recent development in cabin baggage screening, it has been operational for screening hold baggage since many years (Belyavin, 2014; Lazarick, 2001). An airport can use one CIP system to employ a mix of LCBS and RCBS, allowing screeners to use the same user interface regardless of whether they work locally or remotely. Fig. 1 illustrates the two work settings, LCBS and RCBS.

1.3. Effects of noise and social stress

Conveyor belts with trays and baggage, chatting passengers, public announcements, and other sources make checkpoints noisy. Based on field studies at several US airports, Ihrig (2017) reports ambient noise levels in LCBS between 56 and 68 dBA, with an average noise level of 63 dBA (measured without nearby announcements and transient noise). The noise levels in RCBS are much lower, because screeners work in a remote, quiet, office-like environment separated from the busy checkpoints (Kuhn, 2017; Wetter, 2013). The noise levels in RCBS are typically around 40 dBA, comparable to the noise levels in offices (e.g. Delle Macchie et al., 2018). Moreover, whereas screeners in LCBS often experience social stress from difficult and impatient passengers (Bolfing and Schwaninger, 2009; Michel et al., 2014), they have no such exposure in RCBS. Using different tasks, previous studies found that environmental noise and social stress can affect human performance. Several studies have shown that noise can affect cognitive performance (Banbury and Berry, 2005; Boggs and Simon, 1968; Broadbent, 1954; Dalton and Behm, 2007; Hygge and Knez, 2001; Smith, 1988; Taylor et al., 2004; Warner and Heimstra, 1971). Moreover, airport security screeners reported that noise affects baggage screening (Michel et al., 2014); and in a simulated baggage screening task, Yu et al. (2015) found that background speech slowed down processing times. Social stress can also lower performance by reducing the cognitive and emotional resources that can be committed to a task (Hershcovis and Barling, 2010). Furthermore, direct exposure to impatient passengers waiting for their baggage to be checked might induce time pressure, and time pressure has been shown to impact visual search performance (Fleck et al., 2010; Rieger and Manzey, 2022). Performance decrements under high time pressure have also been observed in a simulated X-ray baggage screening task (Rieger et al., 2021). Moreover, evidence from field research indicates that long passenger queues cause screeners to speed up their inspections (Marin et al., 2007). In summary, several studies have found that noise and social stress can negatively affect performance in different tasks. However, most of these studies were based on laboratory experiments with students, and no previous study has compared LCBS and RCBS with airport security screeners in the field.

1.4. Present study

Comparing screening performance in LCBS and RCBS is theoretically interesting and practically relevant because of the increasing interest in RCBS in airports in recent years (Buser and Merks, 2020; Kuhn, 2017). In this study, we analysed a large sample of screener data from a European airport with LCBS and RCBS work settings. We compared screening performance in terms of detecting prohibited articles (hit rate) and the time required to analyse the X-ray images (processing time). The processing time may provide additional information on the mechanisms underlying the potential differences in hit rate. We expected that the less distracting environment in the RCBS and the absence of confrontations with queuing passengers and socially induced time pressure would result in higher hit rates in RCBS than in LCBS. Previous studies have found that time on task (Basner et al., 2008; Buser et al., 2020, 2023; Ghylin et al., 2007; Meuter and Lacherez, 2016) and task load (Buser

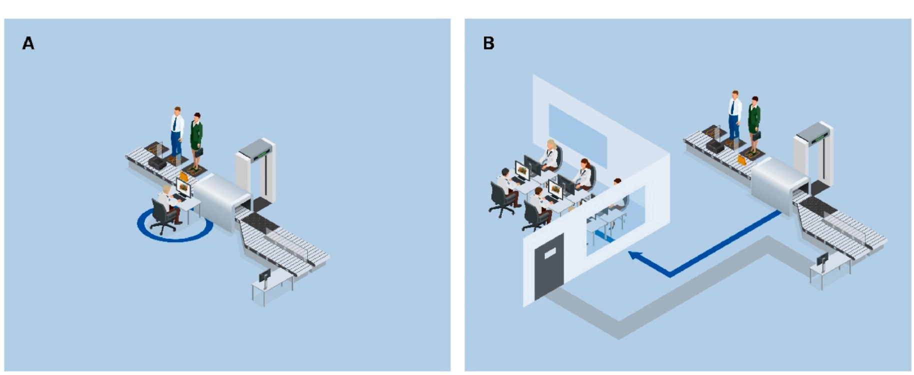

Fig. 1. Work settings in LCBS and RCBS.
Note. Illustrations of (A) local cabin baggage screening (LCBS) and (B) remote cabin baggage screening (RCBS). In LCBS, screeners analyse X-ray images of passenger bags at checkpoints where they are exposed to noise and social stress from passengers. In RCBS, screeners work in remote, quiet, office-like environments separated from the checkpoints.

et al., 2023; Meuter and Lacherez, 2016; Rieger et al., 2021) can reduce screening performance, and that screening performance decreases more with time on task when task load is high (Buser et al., 2023; Meuter and Lacherez, 2016). We expected a similar pattern in our study. However, although negative effects of time on task and task load have been found in both LCBS (Meuter and Lacherez, 2016) and RCBS (Buser et al., 2023), it is unclear whether these effects are equally pronounced in both work settings. Therefore, we analysed whether the effects of time on task and task load differed between LCBS and RCBS.

2. Method

2.1. Sample

Analyses were based on performance data from 1482 screeners who had worked at a European airport over two years. The same screeners worked in both work settings, analysing images in LCBS and RCBS. The same CIP system (with 2D multi-view X-ray imaging technology) was used for both work settings. It logged the screener identity, analyst station location, and screening performance. Due to the CIP system used at the airport, false alarm rates and processing times for target-absent images (i.e., images of bags that do not contain prohibited articles) were unavailable. However, we analysed the screening performance on so called threat image projection images, which is explained in the next section. The CIP data showed only the total of number of processed X-ray images per screener at the day level and did not distinguish between work settings. The task load could therefore only be traced back to a specific work setting for days when screeners had worked in either LCBS or RCBS. We only included such days in the analysis (1706 screeners and 669,168 TIP events representing 69.6% of all TIP data).

2.2. Screening task and threat image projection

Work at the security checkpoint is typically organised in teams, with members rotating through different positions during their shift. In addition to the X-ray image analysis of passenger baggage, work rotation typically includes other positions such as the manual search of passengers or helping passengers divest their belongings. Screeners screen cabin baggage for up to approximately 20 min before proceeding to the next position and usually complete several screening sessions (periods during which a screener is actively working at an X-ray screening position) per work shift. Screeners rarely encounter threat items such as bombs and guns during routine baggage screening. Because the detection of targets can become worse when they are rare (Godwin et al., 2010; Wolfe et al., 2005, 2007), many airports use threat image projection (TIP) to artificially increase screeners' exposure to threat items. TIP systems project pre-recorded X-ray images of threat items (fictional threat images, FTIs) onto images of a random selection of passenger bags screened at the checkpoint (Cutler and Paddock, 2009; Hofer and Schwaninger, 2005; Schwaninger, 2006; Skorupski and Uchroński, 2016). Therefore, a predefined percentage of the images inspected by screeners contain a projected FTI (Fig. 2). If the screener suspects a threat in an X-ray image, they indicate this by pressing a button or marking the threat's location. If a suspected threat is a projected FTI, the TIP system informs the screener of the successful detection. The screener is also informed if they miss an FTI. For each TIP image, the system records the time it was analysed, whether it was detected or missed, and how long the image was displayed before the screener indicated their decision. Therefore, TIP can be used not only to tackle rarely occurring threats during daily screening routines but also to assess the individual performance of screeners (Buser et al., 2023; Cutler and Paddock, 2009; Hofer and Schwaninger, 2005; Meuter and Lacherez, 2016).

2.3. Design

We compared the hit rate and processing time as dependent variables by analysing TIP data from an international airport with both work settings. We included the work setting (RCBS vs LCBS), time on task, and task load as independent variables in our analyses. We examined main effects and interactions. Because screening performance might depend on baggage characteristics that potentially change with the season, we included season as a control variable in our analyses.

2.4. Measures

The hit rate was defined as the percentage of detected FTIs. The processing time was defined as the time between the appearance of an X-ray image on the screen and when screeners pressed the respective button to indicate their decision on whether the image contained a prohibited article. Processing times were examined separately, depending on whether screeners detected a target (PT_{hit}) or missed it (PT_{miss}). They were log-transformed to normalise the residuals and standardised (using a z transformation) before model fitting (Weisberg, 2005). The variable screening location indicates whether the image was analysed while the screener was working either in the lane (i.e., in LCBS) or remotely (i.e., in RCBS). Time on task was calculated as the time elapsed since the first event of the session. Furthermore, task load was calculated as the total number of bags analysed by the screener on that day divided by the total screening duration on the same day. To calculate the latter, we computed the duration of each session as the time from the first reported event (i.e., hit, miss, or non-TIP alarm) within a session to the last reported event within the same session. Time on task and task load were standardised (following a z transformation) prior to model fitting. To control for seasonal effects on the TIP hit rate owing to seasonal differences in baggage characteristics, the variable season was included in the analyses by dividing the year into six 2-month bins (starting in January and February).

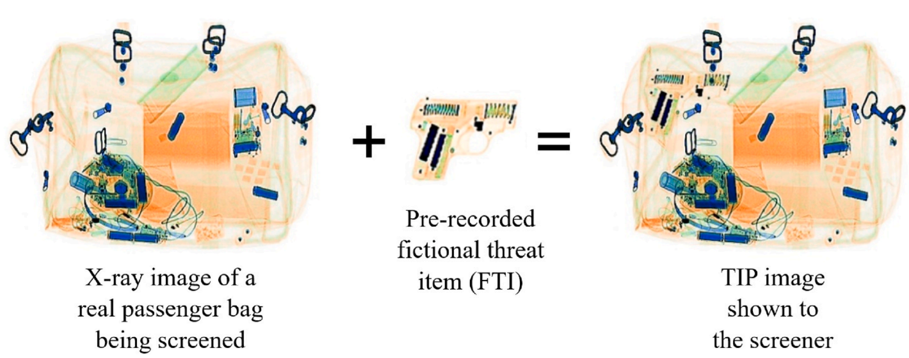

Fig. 2. Threat image projection (TIP) in cabin baggage screening.
Note. TIPs are X-ray images of real passenger bags containing a projection of a pre-recorded fictional threat item (FTI). Adapted from Schwaninger (2006).

Table 1
Compared models for hit rate and processing times.

Model 1	Screening location + Time on task + Task load + Season + (1|Screener) + (1|Session)
Model 2	Model 1 + Time on task × Task load
Model 3	Model 1 + Screening location × Time on task + Screening location × Task load
Model 4	Model 2 + Screening location × Time on task + Screening location × Task load + Screening location × Time on task × Task load

Note. Model specifications tested during model comparisons for hit rate and processing times.

2.5. Exclusions

On rare occasions, the task load was unreasonably high, reaching up to 86 bags per minute. This may have been due to data-logging problems in the CIP system. Days with a task load of more than 30 images/min suggested improbably low average processing times of less than 2 s per image. They were thus excluded from the analyses (excluding 1987 decisions, ~0.3% of the data). For all analyses, decisions were excluded for screeners who worked in only one of the two work settings over the analysed period (963 decisions, ~0.1% of the data). Airports usually analyse TIP performance only for screeners who have analysed at least 20 TIP images over six months. Therefore, we excluded screeners with an average of fewer than 20 decisions on TIP images in total or fewer than 40 per year during the observation period (excluding 991 decisions, ~0.1% of the data). To remove irregular sessions, sessions shorter than 5 min (e.g., due to logins for system testing) or longer than 25 min (e.g., because screeners did not log out properly after screening) were excluded (excluding 21,122 decisions, ~3.2% of the data). Because screening activity was very low at night, we excluded images recorded during night shifts between 10 p.m. and 3 a.m. (excluding 2070 decisions, ~0.3% of the data). In total, 25,063 decisions (3.8%) were excluded from the final analysis. The final dataset used for the model calculations contained the decisions of 1482 screeners made on 642,035 TIP images; 58% were analysed in the lane (LCBS) and 42% were analysed in the remote room (RCBS).

2.6. Analyses

Due to the data's hierarchical structure (decisions nested within sessions that are nested within screeners), data were analysed using mixed models (Snijders and Bosker, 2012). All models included the user and session as random factors. Models were fitted using the lmer and glmer functions of the lme 4 package (Kuznetsova et al., 2017) in R (R Core Team, 2020). Reported p values were computed using the Satterthwaite formula for degrees of freedom. The hit rate was analysed by fitting a logistic mixed model using the maximum likelihood [ML] and Nelder–Mead optimisation methods. The processing times were modelled by fitting linear mixed models using restricted maximum likelihood [REML] and nloptwrap optimisers. We analysed PT_{hit} and PT_{miss} separately because positive (hits) and negative responses on target-present images (misses) are considered to follow distinct search dynamics. Positive responses are typically faster than negative responses (Koller et al., 2008; Rieger et al., 2021; Wolfe, 1998; Wolfe and Van Wert, 2010). To select the appropriate model for each dependent variable (hit rate, PT_{hit}, and PT_{miss}), we first tested a model with only the main effects against a model that also included the interaction time on task × task load. An interaction was included when the likelihood ratio test indicated improved model fit. We then tested either the main effect or interaction model against a model that also included the interactions of screening location with all other independent variables in the model but not the control variable season. Table 1 presents the models we compared to each other.

3. Results

For the hit rate and PT_{miss}, the interaction time on task × task load improved the model fit. However, including interactions with screening location increased the model fit only for PT_{miss}. For PT_{hit}, the interaction time on task × task load did not improve the model fit. However, the model fit increased when interactions with screening location were included. Table 2 presents the model comparisons and final models for each dependent variable. Fig. 3 shows the marginal effects of time on task on the hit rate (top), PT_{hit} (bottom left), and PT_{miss} (bottom right), depending on screening location and task load. For readability, estimates and confidence intervals were back-transformed to absolute rates and processing times without any bias correction. The logistic mixed model for the hit rate revealed a significant main effect of screening location ($b = 0.104$, $SE = 0.010$, $p < .001$), indicating that the hit rate was higher for the RCBS than for the LCBS. Main effects for time on task ($b = -0.025$, $SE = 0.004$, $p < .001$) and task load ($b = -0.019$, $SE = 0.004$, $p < .001$) were also significant, indicating that hit rate decreased with increasing time on task and task load. Moreover, a significant task load × time on task interaction was found ($b = -0.009$, $SE = 0.004$, $p = .017$), indicating that the hit rate decreased more strongly with time on task when task load was high compared to when task load was lower. Table 3 shows the estimated odds ratios and confidence intervals for all variables in the model. For PT_{hit}, there was a significant main effect of screening location ($b = -0.070$, $SE = 0.003$, $p < .001$). Moreover, there was a significant main effect of task load ($b = -0.047$, $SE = 0.002$, $p < .001$) and a significant task load × screening location interaction ($b = -0.027$, $SE = 0.003$, $p < .001$). This indicated that PT_{hit} decreased more strongly with increasing task load in the RCBS than in the LCBS. The main effect of time on task ($b = 0.003$, $SE = 0.002$, $p = .053$) was not significant. However, a significant time on task × screening location interaction was found ($b = -0.007$, $SE = 0.002$, $p = .004$), whereby PT_{hit} decreased with time on task in the RCBS and not in the LCBS (Fig. 3). The regression outputs and confidence intervals of the model on PT_{hit} are listed in Table 4 (left). For PT_{miss}, there were significant main effects of screening location ($b = 0.088$, $SE = 0.008$, $p < .001$), time on task ($b = -0.045$, $SE = 0.004$, $p < .001$), and task load ($b = -0.122$, $SE = 0.004$, $p < .001$). The interaction time on task × task load was also significant ($b = -0.018$, $SE = 0.003$, $p < .001$). This indicated that PT_{miss} decreased more strongly with time on task when the task load was high. Moreover, the interactions for time on task × screening location ($b = -0.015$, $SE = 0.007$, $p = .022$) and task load × screening location ($b = -0.018$, $SE = 0.008$, $p = .028$) were significant, indicating stronger negative effects of time on task and task load in the RCBS than in the LCBS. A significant three-way interaction for time on task × task load × screening location was found ($b = -0.020$, $SE = 0.008$, $p = .010$). This indicated that the negative effect of time on task was amplified more strongly by task load in the RCBS than in the LCBS. Table 4 (right) lists the regression outputs and confidence intervals of the model on PT_{miss}.

4. Discussion

With RCBS, airport security screeners can work in a quiet, office-like environment away from busy checkpoints (LCBS) where screeners are exposed to noise and social stress from passengers (Bolfing and Schwaninger, 2009; Buser and Merks, 2020; Kuhn, 2017; Michel et al., 2014). The primary aim of this study was to compare screening performance in LCBS and RCBS. To this end, we analysed extensive TIP data from an international European airport. We further investigated whether the effects of time on task and task load differed between the two work settings. We found a higher hit rate and longer processing time

Table 2
Model comparisons and final models for hit rate and processing times.

		AIC	Loglik	Deviance	χ^2	Df	p
Test whether including the interaction time on task × task load improved model fit							
Hit rate	Model 1	460427	−230202	460405	–	–	–
	Model 2	460423	−230200	460399	5.410	1	.020*
PT_{hit}	Model 1	1475011	−737494	1474987	–	–	–
	Model 2	1475012	−737493	1474986	0.765	1	.382
PT_{miss}	Model 1	193081	−96528	193057	–	–	–
	Model 2	193037	−96505	193011	45.856	1	<.001***
Test whether including interactions with screening location improved model fit							
Hit rate	**Model 2**	460423	−230200	460399	–	-	-
	Model 4	460426	−230198	460396	3.199	3	.362
PT_{hit}	Model 1	1475011	−737494	1474987	–	–	–
	Model 3	1474942	−737457	1474914	72.759	2	<.001***
PT_{miss}	Model 2	193037	−96505	193011	–	–	–
	Model 4	193030	−96499	192998	13.168	3	.004**

Note. PT_{hit} = Processing time for hits; PT_{miss} = Processing time for misses; AIC = Akaike's information criterion; Loglik = log likelihood; χ^2 = chi-square test statistic; Df = difference in degrees of freedom of the compared models. The final model for each dependent variable is shown in bold. *$p < .05$. **$p < .01$. $p^{***} < .001$.

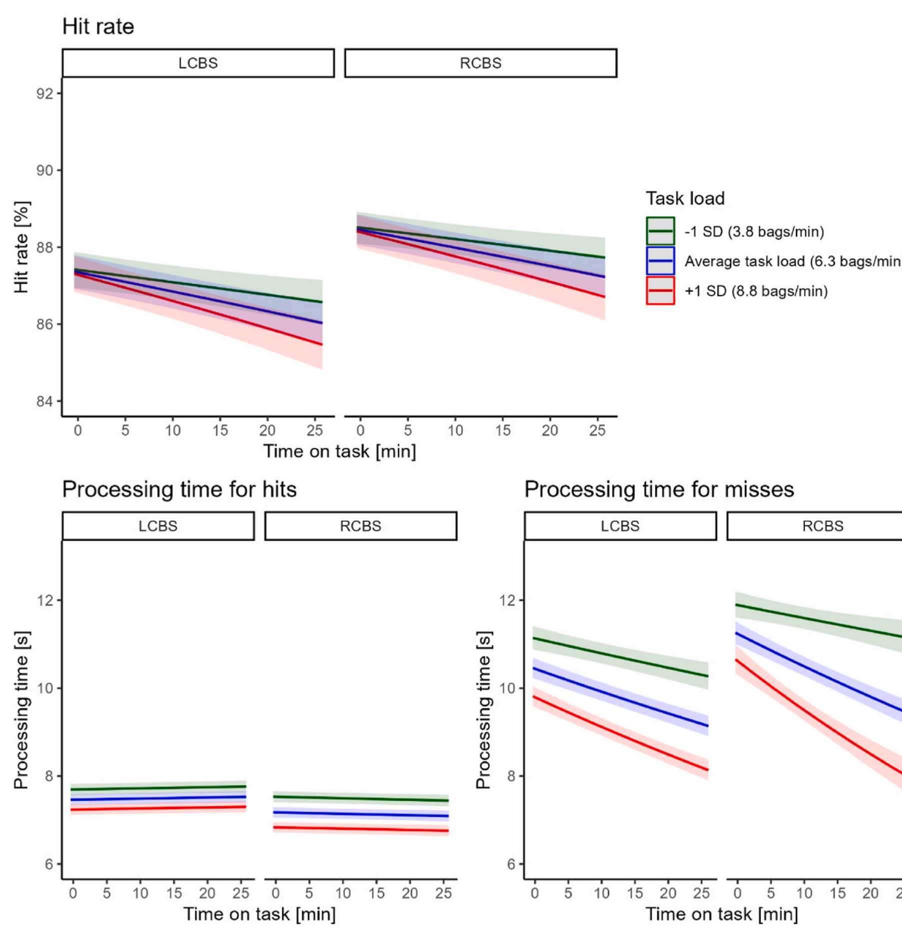

Fig. 3. Marginal effects of time on task on hit rate and processing times for hits and misses.
Note. Hit rate (top) and processing times (bottom) are shown based on model predictions for screening location, time on task, and three different levels of task load (average task load, ±1 *SD*).

in RCBS compared to LCBS. In both working settings, the hit rate and processing times for misses decreased with time on task, whereby the decrease in the hit rate was stronger when task load was high.

4.1. Remote cabin baggage screening versus local cabin baggage screening

At airport security checkpoints, various factors contribute to noise levels, including conveyor belts with trays and baggage, passenger conversations, and public announcements. In LCBS, background noise levels of 56–68 dBA have been reported, dependent on the size and architecture of the security checkpoint (Ihrig, 2017). However, in RCBS, screeners operate in a quieter, office-like environment away from the bustling checkpoints. Here, noise levels are significantly lower (Kuhn, 2017; Wetter, 2013), usually around 40 dBA, comparable to typical office noise levels (Delle Macchie et al., 2018). Additional measurements that we conducted at an international European airport confirmed substantial differences regarding the average noise levels in the two work settings (we measured an average of 65 dBA in LCBS and 32 dBA in

Table 3
Fixed effects, random effects, and variance explanation of the mixed model for the hit rate.

Predictor	Hit rate		
	Odds ratios	95% CI	*p*
(Intercept)	6.652	[6.403, 6.911]	< .001***
Screening location [RCBS]	1.109	[0.984, 0.998]	< .001***
Time on task	0.974	[0.967, 0.982]	< .001***
Task load	0.981	[0.973, 0.990]	< .001***
Time on task × Task load	0.991	[0.984, 0.998]	.017*
Season [2]	0.978	[0.949, 1.007]	.131
Season [3]	0.923	[0.896, 0.951]	< .001***
Season [4]	0.917	[0.891, 0.944]	< .001***
Season [5]	0.988	[0.959, 1.017]	.402
Season [6]	1.101	[1.087, 1.132]	< .001***
Random effects			
σ^2	3.29		
τ^{00} Session	.08		
τ^{00} User	.28		
ICC	.10		
N User	1482		
N Session	334,116		
Observations	642,035		
Marginal R^2/conditional R^2	.002/.101		

Note. σ^2 = residual variance or within-subject variance; τ^{00} = random intercept variance or between-subject variance; marginal R^2 = variance explanation through the fixed effects; conditional R^2 = variance explanation through the fixed and random effects; ICC = intraclass correlation. *p < .05. ***p < .001.

RCBS). In our study, as expected, screeners showed a higher detection rate of prohibited articles (hit rate) in RCBS than in LCBS. This is consistent with Michel et al. (2014) who reported that screeners perceive noise as relevant to baggage-screening performance. Moreover, several studies using different tasks, mostly in the laboratory with students, have shown that environmental noise can be distractive and can lower cognitive performance on various tasks (e.g., Banbury and Berry, 2005; Dalton and Behm, 2007; Hygge and Knez, 2001; Smith, 1988; see Szalma and Hancock, 2011, for a comprehensive review). Studies on visual search and signal detection under noisy conditions imply that ambient noise can be especially relevant for search performance when task complexity is high (Boggs and Simon, 1968; Broadbent, 1954; Warner and Heimstra, 1971). However, the effects of noise on performance have also been observed at relatively low levels of task complexity. For instance, Taylor et al. (2004) found that intermittent and random noise lowered performance in a visual search task and concluded that inspectors could be distracted by intermittent or random noise patterns at the workplace even when conducting a relatively easy search task. Yu et al. (2015) investigated the effects of background speech on performance in a simulated X-ray baggage screening task. They found higher processing times with background speech than in a non-speech condition. The effect on the hit rate was not significant, possibly due to the sample size that would not allow the detection of small effects. At the checkpoint, environmental noise includes background speech as well as other sources of noise such as baggage handling and the operation of screening equipment.

Another plausible contributor to the higher hit rate in RCBS is that screeners are not directly exposed to queuing passengers, whereas screeners in LCBS are regularly confronted with difficult passengers (Bolfing and Schwaninger, 2009; Michel et al., 2014). When confronted with customers who are rude, impatient, or otherwise ungrateful, emotional regulation becomes necessary and involves emotional labour (Chu et al., 2012; Grandey et al., 2002; Grebner et al., 2003; Hochschild, 2022; Morris and Feldman, 1996). In such stressful interactions, task performance can be impaired because cognitive and emotional resources are strained and cannot be directed towards the task (Hershcovis and Barling, 2010). Additionally, passengers wait impatiently in the security lane for their baggage to be checked. This can lead to stress due to socially induced time pressure, because screeners can feel urged to hurry to reduce passenger waiting times. Studies have shown that time pressure reduces visual search performance (Fleck et al., 2010; Rieger et al., 2021; Rieger and Manzey, 2022). Additionally, Marin et al. (2007) found that screeners accelerated their inspection when passenger queues were long.

The differences in hit rate and processing times found in our study cannot be explained by a speed–accuracy trade-off (McCarley, 2009; Wickelgren, 1977) because hits were processed faster and not slower in RCBS than in LCBS. Instead, the lower socially induced time pressure in RCBS might have led to changes in the perceived payoff for decisions on whether a passenger bag contains a prohibited article. Changes in

Table 4
Fixed effects, random effects, and variance explanation of the mixed model for PT_{hit} and PT_{miss}.

Predictors	PT_{hit}			PT_{miss}		
	Estimates	95% CI	*p*	Estimates	95% CI	*p*
(Intercept)	0.252	[0.227, 0.276]	< .001***	0.142	[0.111, 0.173]	< .001***
Screening location [RCBS]	−0.070	[-0.077, −0.063]	< .001***	0.088	[0.072, 0.104]	< .001***
Time on task	0.003	[-0.000, 0.006]	.053	−0.045	[-0.052, −0.037]	< .001***
Task load	−0.047	[-0.050, −0.043]	< .001***	−0.122	[-0.129, −0.114]	< .001***
Time on task × Task load				−0.018	[-0.024, −0.011]	< .001***
Time on task × Screening location [RCBS]	−0.007	[-0.012, −0.002]	.004**	−0.015	[-0.028, −0.002]	.022*
Task load × Screening location [RCBS]	−0.027	[-0.034, −0.021]	< .001***	−0.018	[-0.034, −0.002]	.028*
Time on task × Task load × Screening location [RCBS]				−0.020	[-0.035, −0.005]	.009**
Season [2]	0.017	[0.008, 0.027]	< .001***	0.023	[-0.000, 0.046]	.053
Season [3]	0.026	[0.016, 0.035]	< .001***	−0.002	[-0.025, 0.021]	.887
Season [4]	0.021	[0.011, 0.030]	< .001***	−0.024	[-0.047, −0.002]	.033*
Season [5]	−0.008	[-0.018, 0.001]	.090	−0.025	[-0.047, −0.002]	.030*
Season [6]	−0.048	[-0.057, −0.038]	< .001***	−0.008	[-0.031, 0.016]	.513
Random effects						
σ^2	0.66			0.60		
τ^{00}	.14 Session			0.08 Session		
	.21 User			0.21 User		
ICC	.34			0.33		
N	1482 User			1475 User		
	313,864 Session			68,971 Session		
Observations	564,684			77,351		
Marginal R^2/conditional R^2	.004/.346			.026/.346		

Note. σ^2 = residual variance or within-subject variance; τ^{00} = random intercept variance or between-subject variance; marginal R^2 = variance explanation through the fixed effects; conditional R^2 = variance explanation through the fixed and random effects; ICC = intraclass correlation.*$p < .05$. **$p < .01$. ***$p < .001$

perceived decision payoffs typically lead to a shift in people's decision criteria (Green and Swets, 1966; Macmillan and Creelman, 2005; Sterchi et al., 2019). In RCBS, screeners could decide more often that a prohibited article is present (i.e., apply a more liberal criterion) because the immediate consequences of producing a false alarm (i.e., sending a harmless bag to a secondary search, thereby lowering passenger throughput) are less salient. When such a criterion shift occurs, hits become more frequent and faster, whereas "target-absent" responses on target-absent images (i.e., images not containing a prohibited article) become slower (Wolfe et al., 2007; Wolfe and Van Wert, 2010).

In our study, we could measure decisions and processing times only for target-present images (i.e., TIP images) but not for target-absent images. However, other studies have found that response times for negative decisions on target-present images (i.e., misses) follow a similar pattern to that of negative decisions on target-absent images. Therefore, the response pattern expected under a criterion shift is consistent with that observed in our study: a higher hit rate in RCBS associated with faster processing times for hits and slower processing times for misses compared with LCBS.

4.2. Time on task and task load

As expected, the hit rate decreased with increasing time on task and task load in both settings, with a stronger decrease over time when task load was high. This result is consistent with previous studies reporting that the hit rate in X-ray baggage screening decreases with time on task (Basner et al., 2008; Buser et al., 2023; Ghylin et al., 2007; Meuter and Lacherez, 2016) and that the performance decrement is stronger when task load is high (Buser et al., 2023; Meuter and Lacherez, 2016). We observed a relatively small decrease in the hit rate over time when task load was low. For higher task loads, the decrease over time was stronger. In addition to the decrease in the hit rate, we observed a decrease in the processing time for misses with increasing time on task. However, the processing time for hits hardly changed with increasing task time. This suggests that the observed effects of time on task are not the result of a mere speed–accuracy trade-off. However, the observed changes in processing time with time on task also did not show the pattern expected for a mere criterion shift. Moreover, increasingly faster responses for misses contradict theories that assume that a depletion of attentional resources over time is the source of a drop in performance (Helton and Warm, 2008; Matthews et al., 2010).

A plausible explanation for the decrease in hit rate and processing time with time on task comes from dynamic resource theory (DART; Rubinstein, 2020). DART assumes that individuals apply implicit strategies to prevent the potential loss of cognitive resources over time. To preserve their cognitive resources, they invest less time in their searches. Lower processing times then lead to fewer target-present responses (Rubinstein, 2020). DART could also explain the interaction between time on task and task load, with hit rate and processing time of misses decreasing more rapidly when the task load is high. With a high task load demanding more cognitive resources, adapting behaviour to prevent the loss of these resources becomes more urgent. DART provides a framework to explain the reduced hit rate, false alarm rate, and response time typically observed in search tasks with rare targets (Basner et al., 2008; Drury, 2015; Thomson et al., 2014). Further research could help to explain why the decrease in processing time in our study was driven mainly by images in which individuals did not detect a target (PT_{miss}).

Interestingly, we did not find differences between RCBS and LCBS regarding the effects of time on task and task load on the hit rate. This implies that even a relatively sheltered and quiet work environment in RCBS does not alleviate the decline in performance over time. For the processing time of misses, however, we found a three-way interaction between screening location, time on task, and task load: the speed was accelerated more strongly by task load in RCBS than in LCBS. Thus, whereas hit rates remained higher in RCBS over the whole session, processing time for misses converged towards processing time in LCBS when the task load was high. This may have been because the screeners started off more slowly in RCBS and had more scope for speeding up.

4.3. Practical implications

Our results suggest that the work environment can significantly impact screening performance. The estimated increase in the hit rate for RCBS was 1.14 percentage points. Given that, on average, 13.07% of the TIP images were missed in LCBS, the increase in the hit rate implies 8.7% fewer misses in RCBS. This roughly means that for every 10 prohibited items that are missed in LCBS, one is found in RCBS. Simultaneously, higher processing times for misses in RCBS suggest that screeners needed approximately 0.6 s more time for image analysis when working remotely. Higher hit rates and longer processing times in RCBS than in LCBS indicate a trade-off between performance and throughput, which airports need to consider and weigh individually. Notably, the effects of the work setting found in our study were relatively small compared with the differences between individual screeners. This is consistent with previous studies showing that screening performance depends strongly on individual factors such as the screener's cognitive abilities (Hardmeier et al., 2005; Hättenschwiler et al., 2019; Peltier and Becker, 2017, 2020) along with the quality and amount of computer-based training and the higher X-ray image interpretation competence that this can produce (Halbherr et al., 2013; Koller et al., 2008; McCarley et al., 2004; Schuster et al., 2013). Focusing on the optimal selection and training of screeners promises a greater improvement in the hit rate than introducing RCBS. However, RCBS can offer additional improvements.

Independent of the work environment, we found that the hit rate decreased with time on task; moreover, this decrease was steeper when task load was high. That is, time on task was more relevant under high task load. Interestingly, the decrease in hit rate with time on task was equally pronounced in both work settings. However, aligned with recent findings from Buser et al. (2023), we found that the changes in screening performance with time on task that occurred within typical session lengths (approximately 20 min) were relatively small compared to the inter-individual differences between screeners.

4.4. Limitations

Based on the available TIP data, we could not calculate the false alarm rate (i.e., the percentage of bags that were harmless but wrongly classified as containing a prohibited article) and the processing time for target-absent images (i.e., processing times for images of bags that did not contain a prohibited article). This represents a limitation often encountered when analysing field data from TIP systems (Meuter and Lacherez, 2016; Skorupski and Uchroński, 2016). Data on false alarm rates are necessary to make definitive conclusions about whether the observed changes in hit rates are due to a shift in screener response bias or changes in sensitivity. The processing time for target-absent images could contribute to gaining a more detailed understanding of the dynamics underlying the observed effects on the hit rate. The processing time for target-absent images is also relevant from a practical perspective because most images in the daily screening routine do not contain a prohibited article, thus driving the overall throughput of screened passenger baggage at the checkpoint. Although the average noise level and social stress might be the most notable differences between the two work settings, they can also differ in other aspects that might affect performance (e.g. lighting). Therefore, our study cannot attribute the observed performance differences directly to noise and social stress alone. Furthermore, the screeners in our study alternated between working in LCBS and RCBS, and our analyses focused on performance differences between the two work settings. However, there might be potential long-term effects of working in RCBS (e.g., breaks from being exposed to a noisy environment in LCBS) that our analyses failed to capture. Furthermore, our study was limited to one specific airport and may not be generalisable to airports with different levels of stressors or screener populations. Survey data or structured interviews could extend our research by identifying the stressors that foster or impede high performance in baggage screening and thus provide advice on future checkpoint architectures.

4.5. Conclusion

In RCBS, screeners work in a remote, quiet, office-like environment away from busy checkpoints where screeners are exposed to noise and social stress from passengers. Our study provides the first evidence that RCBS results in better detection of prohibited articles than LCBS at the expense of longer processing times that cannot be explained by a speed-accuracy trade-off. Furthermore, our results suggest that decrements in the hit rate over time on task are similar for both work settings and are accelerated when task load is high. Importantly, the effects of work setting, time on task, and task load were relatively small compared with the differences between individual screeners.

CRediT authorship contribution statement

Marius Latscha: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Adrian Schwaninger:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Jürgen Sauer:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Yanik Sterchi:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

This study was supported by the Swiss National Science Foundation [grant number: 100019_188808].

References

Banbury, S.P., Berry, D.C., 2005. Office noise and employee concentration: identifying causes of disruption and potential improvements. Ergonomics 48 (1), 25–37. https://doi.org/10.1080/00140130412331311390.

Basner, M., Rubinstein, J., Fomberstein, K.M., Coble, M.C., Ecker, A., Avinash, D., Dinges, D.F., 2008. Effects of night work, sleep loss and time on task on simulated threat detection performance. Sleep 31 (9), 1251–1259. https://doi.org/10.5665/sleep/31.9.1251.

Belyavin, A., 2014. Simulating the impact of remote screening on search comb capacity. In: International Carnahan Conference on Security Technology, pp. 1–6. https://doi.org/10.1109/CCST.2014.6987004.

Biggs, A.T., Mitroff, S.R., 2015. Improving the efficacy of security screening tasks: a review of visual search challenges and ways to mitigate their adverse effects. Appl. Cognit. Psychol. 29, 142–148. https://doi.org/10.1002/acp.3083.

Boggs, D.H., Simon, J.R., 1968. Differential effect of noise on tasks of varying complexity. J. Appl. Psychol. 52 (2), 148–153. https://doi.org/10.1037/h0025496.

Bolfing, A., Halbherr, T., Schwaninger, A., 2008. How image based factors and human factors contribute to threat detection performance in X-ray aviation security screening. In: Holzinger, A. (Ed.), HCI and Usability for Education and Work. USAB 2008, Lecture Notes in Computer Science, vol. 5298. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-89350-9_30.

Bolfing, A., Schwaninger, A., 2009. Selection and pre-employment assessment in aviation security X-ray screening. In: 43rd Annual 2009 International Carnahan Conference on Security Technology, pp. 5–12. https://doi.org/10.1109/CCST.2009.5335571.

Broadbent, D.E., 1954. Some effects of noise on visual performance. Q. J. Exp. Psychol. 6 (1), 1–5. https://doi.org/10.1080/17470215408416643.

Buser, D., Merks, S., 2020. Centralised image processing: challenges, trends and time on task. Aviat Secur. Int. 33–35. In: https://www.researchgate.net/publication/352778346_Centralised_Image_Processing_challenges_trends_and_time_on_task&_iepl=. (Accessed 4 October 2023).

Buser, D., Schwaninger, A., Sauer, J., Sterchi, Y., 2023. Time on task and task load in visual inspection: a four-month field study with X-ray baggage screeners. Appl. Ergon. 111, 103995 https://doi.org/10.1016/j.apergo.2023.103995.

Buser, D., Sterchi, Y., Schwaninger, A., 2020. Why stop after 20 minutes? Breaks and target prevalence in a 60-minute X-ray baggage screening task. Int. J. Ind. Ergon. 76, 102897 https://doi.org/10.1016/j.ergon.2019.102897.

Chu, K.H., Baker, M.A., Murrmann, S.K., 2012. When we are onstage, we smile: the effects of emotional labor on employee work outcomes. Int. J. Hospit. Manag. 31 (3), 906–915. https://doi.org/10.1016/J.IJHM.2011.10.009.

Cutler, V., Paddock, S., 2009. Use of threat image projection (TIP) to enhance security performance. In: 43rd Annual International Carnahan Conference on Security Technology, pp. 46–51. https://doi.org/10.1109/CCST.2009.5335565.

Dalton, B.H., Behm, D.G., 2007. Effects of noise and music on human and task performance: a systematic review. Occup. Ergon. 7 (3), 143–152. https://doi.org/10.3233/OER-2007-730.

Delle Macchie, S., Secchi, S., Cellai, G., 2018. Acoustic issues in open plan offices: a typological analysis. Buildings 8 (11), 161. https://doi.org/10.3390/BUILDINGS8110161, 2018, Vol. 8, Page 161.

Donnelly, N., Muhl-Richardson, A., Godwin, H.J., Cave, K.R., 2019. Using eye movements to understand how security screeners search for threats in X-ray baggage. Vision 3, 24. https://doi.org/10.3390/VISION3020024, 2019, Vol. 3, Page 24.

Dorton, S.L., Liu, D., 2016. Effects of baggage volume and alarm rate on airport security screening checkpoint efficiency using queuing networks and discrete event simulation. Hum. Factors Ergon. Manuf. 26 (1), 95–109. https://doi.org/10.1002/HFM.20616.

Drury, C.G., 2015. Sustained attention in operational settings. In: Hoffmann, R.R., Hancock, P.A., Scerbo, M.W., Parasuraman, R., Szalma, J.L. (Eds.), The Cambridge Handbook of Applied Perception Research. Cambridge University Press, pp. 769–792. https://doi.org/10.1017/CBO9780511973017.047.

Fleck, M.S., Samei, E., Mitroff, S.R., 2010. Generalized "satisfaction of search": adverse influences on dual-target search accuracy. J. Exp. Psychol. Appl. 16 (1), 60–71. https://doi.org/10.1037/A0018629.

Ghylin, K.M., Drury, C.G., Batta, R., Lin, L., 2007. Temporal effects in a security inspection task: breakdown of performance components. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 51 (2), 93–97. https://doi.org/10.1177/154193120705100209.

Godwin, H.J., Menneer, T., Cave, K.R., Helman, S., Way, R.L., Donnelly, N., 2010. The impact of relative prevalence on dual-target search for threat items from airport X-ray screening. Acta Psychol. 134 (1), 79–84. https://doi.org/10.1016/J.ACTPSY.2009.12.009.

Grandey, A.A., Tam, A.P., Brauburger, A.L., 2002. Affective states and traits in the workplace: diary and survey data from young workers. Motiv. Emot. 26 (1).

Grebner, S., Semmer, N., Faso, L.L., Gut, S., Kälin, W., Elfering, A., 2003. Working conditions, well-being, and job-related attitudes among call centre agents. European J. Work and Organ. Psychol. 12 (4), 341–365. https://doi.org/10.1080/13594320344000192.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. John Wiley & Sons, Inc., New York. https://psycnet.apa.org/record/1967-02286-000. (Accessed 4 October 2023).

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. https://doi.org/10.1080/10508414.2011.582455.

Hardmeier, D., Hofer, F., Schwaninger, A., 2005. The X-ray object recognition test (X-ray ORT) - a reliable and valid instrument for measuring visual abilities needed in X-ray screening. In: Proceedings 39th Annual 2005 International Carnahan Conference on Security Technology, pp. 189–192. https://doi.org/10.1109/ccst.2005.1594876.

Hättenschwiler, N., Merks, S., Sterchi, Y., Schwaninger, A., 2019. Traditional visual search vs. X-ray image inspection in students and professionals: are the same visual-cognitive abilities needed? Front. Psychol. 10 https://doi.org/10.3389/fpsyg.2019.00525.

Helton, W.S., Warm, J.S., 2008. Signal salience and the mindlessness theory of vigilance. Acta Psychol. 129 (1), 18–25. https://doi.org/10.1016/j.actpsy.2008.04.002.

Hershcovis, M.S., Barling, J., 2010. Towards a multi-foci approach to workplace aggression: a meta-analytic review of outcomes from different perpetrators. J. Organ. Behav. 31 (1), 24–44. https://doi.org/10.1002/job.621.

Hochschild, A.R., 2022. The managed heart. Working in America: Continuity, Conflict, and Change in a New Economic Era, fifth ed., pp. 40–48. https://doi.org/10.4324/9781003199588-7/managed-heart-arlie-russell-hochschild

Hofer, F., Schwaninger, A., 2005. Using threat image projection data for assessing individual screener performance. WIT Trans. Built. Environ. 82, 417–426. https://doi.org/10.2495/safe050411.

Hygge, S., Knez, I., 2001. Effects of noise, heat and indoor lighting on cognitive performance and self-reported affect. J. Environ. Psychol. 21 (3), 291–299. https://doi.org/10.1006/jevp.2001.0222.

Ihrig, W., 2017. Improving intelligibility of airport terminal public address systems. Airport Cooperative Research Programm 175. https://doi.org/10.17226/24839.

International Air Transport Association, 2019. Airport development reference manual. Forecasting and Planning Sections Produced in Collaboration with. ACI eleventh ed. https://www.iata.org/en/publications/store/airport-development-reference-manual/. (Accessed 4 October 2023).

Koller, S.M., Drury, C.G., Schwaninger, A., 2009. Change of search time and non-search time in X-ray baggage screening due to training. Ergonomics 52 (6), 644–656. https://doi.org/10.1080/00140130802526935.

Koller, S.M., Hardmeier, D., Michel, S., Schwaninger, A., 2008. Investigating training, transfer and viewpoint effects resulting from recurrent CBT of X-ray image interpretation. J. Trans. Secur. 1 (2), 81–106. https://doi.org/10.1007/s12198-007-0006-4.

Kuhn, M., 2017. Centralised image processing: the impact on security checkpoints. Aviat Secur. Int. 23 (5), 28–30. https://api.semanticscholar.org/CorpusID:4800163. (Accessed 4 October 2023).

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest package: tests in linear mixed effects models. J. Stat. Software 82 (13), 1–26. https://doi.org/10.18637/jss.v082.i13.

Lazarick, R., 2001. Applications of technology in airport access control. In: Proceedings IEEE 35th Annual International Carnahan Conference on Security Technology, pp. 85–95. https://doi.org/10.1109/ccst.2001.962817.

Macmillan, N.A., Creelman, C.D., 2005. Detection Theory: a User's Guide, second ed. Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey https://doi.org/10.4324/9781410611147.

Marin, C.V., Drury, C.G., Batta, R., Lin, L., 2007. Human factors contributes to queuing theory: Parkinson's law and security screening. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 51 (10), 602–606. https://doi.org/10.1177/154193120705101007.

Matthews, G., Warm, J.S., Reinerman-Jones, L.E., Langheim, L.K., Washburn, D.A., Tripp, L., 2010. Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention. J. Exp. Psychol. Appl. 16 (2), 187–203. https://doi.org/10.1037/a0019572.

McCarley, J.S., 2009. Effects of speed–accuracy instructions on oculomotor scanning and target recognition in a simulated baggage X-ray screening task. Ergonomics 52 (3), 325–333. https://doi.org/10.1080/00140130802376059.

McCarley, J.S., Kramer, A.F., Wickens, C.D., Vidoni, E.D., Boot, W.R., 2004. Visual skills in airport-security screening. Psychol. Sci. 15 (5), 302–306. https://doi.org/10.1111/j.0956-7976.2004.00673.x.

Meuter, R.F.I., Lacherez, P.F., 2016. When and why threats go undetected: impacts of event rate and shift length on threat detection accuracy during airport baggage screening. Hum. Factors 58, 218–228. https://doi.org/10.1177/0018720815616306.

Michel, S., Hättenschwiler, N., Kuhn, M., Strebel, N., Schwaninger, A., 2014. A multi-method approach towards identifying situational factors and their relevance for X-ray screening. In: International Carnahan Conference on Security Technology, pp. 1–6. https://doi.org/10.1109/CCST.2014.6987001.

Morris, J.A., Feldman, D.C., 1996. The dimensions, antecedents, and consequences of emotional labor. Acad. Manag. Rev. 21 (4), 986–1010. https://doi.org/10.5465/amr.1996.9704071861.

Peltier, C., Becker, M.W., 2020. Individual differences predict low prevalence visual search performance and sources of errors: an eye-tracking study. J. Exp. Psychol. Appl. 26, 646–658. https://doi.org/10.1037/XAP0000273.

Peltier, C., Becker, M.W., 2017. Individual differences predict low prevalence visual search performance. Cogn Res Princ Implic 2. https://doi.org/10.1186/S41235-016-0042-3.

R Core Team, 2020. R: a language and environment for statistical computing. https://www.r-project.org/. (Accessed 4 October 2023).

Rieger, T., Heilmann, L., Manzey, D., 2021. Visual search behavior and performance in luggage screening: effects of time pressure, automation aid, and target expectancy. Cogn. Res. 6 (12), 1–12. https://doi.org/10.1186/s41235-021-00280-7.

Rieger, T., Manzey, D., 2022. Understanding the impact of time pressure and automation support in a visual search task. Hum. Factors. https://doi.org/10.1177/00187208221111236.

Rubinstein, J.S., 2020. Divergent response-time patterns in vigilance decrement tasks. J. Exp. Psychol. Hum. Percept. Perform. 46 (10), 1058–1076. https://doi.org/10.1037/xhp0000813.

Schuster, D., Rivera, J., Sellers, B.C., Fiore, S.M., Jentsch, F., 2013. Perceptual training for visual search. Ergonomics 56 (7), 1101–1115. https://doi.org/10.1080/00140139.2013.790481.

Schwaninger, A., 2006. Threat image projection: enhancing performance? Aviat Secur. Int. 36–41. https://www.researchgate.net/publication/259382300_Threat_Image_Projection_enhancing_performance (Accessed 4 October 2023).

Schwaninger, A., Hardmeier, D., Hofer, F., 2005. Aviation security screeners visual abilities & visual knowledge measurement. IEEE Aerospace and Electronic Systems 20 (6), 29–35. https://doi.org/10.1109/MAES.2005.1412124.

Skorupski, J., Uchroński, P., 2016. A human being as a part of the security control system at the airport. Process Eng. 134, 291–300. https://doi.org/10.1016/j.proeng.2016.01.010.

Smith, A.P., 1988. Acute effects of noise exposure: an experimental investigation of the effects of noise and task parameters on cognitive vigilance tasks. Int. Arch. Occup. Environ. Health 60, 307–310. https://doi.org/10.1007/BF00378479.

Snijders, T.A.B., Bosker, R.J., 2012. Multilevel Analysis: an Introduction to Basic and Advanced Multilevel Modeling, second ed. SAGE, Los Angeles https://www.torrossa.com/en/resources/an/4912508. (Accessed 4 October 2023).

Sterchi, Y., Hättenschwiler, N., Schwaninger, A., 2019. Detection measures for visual inspection of X-ray images of passenger baggage. Atten. Percept. Psychophys. 81, 1297–1311. https://doi.org/10.3758/s13414-018-01654-8.

Sterchi, Y., Schwaninger, A., 2015. A first simulation on optimizing EDS for cabin baggage screening regarding throughput. In: Proceedings - International Carnahan Conference on Security Technology 2015-January, pp. 55–60. https://doi.org/10.1109/CCST.2015.7389657.

Szalma, J.L., Hancock, P.A., 2011. Noise effects on human performance: a meta-analytic synthesis. Psychol. Bull. 137 (4), 682–707. https://doi.org/10.1037/a0023987.

Taylor, W., Melloy, B., Dharwada, P., Gramopadhye, A., Toler, J., 2004. The effects of static multiple sources of noise on the visual search component of human inspection. Int. J. Ind. Ergon. 34 (3), 195–207. https://doi.org/10.1016/j.ergon.2004.04.002.

Thomson, D.R., Seli, P., Besner, D., Smilek, D., 2014. On the link between mind wandering and task performance over time. Conscious. Cognit. 27, 14–26. https://doi.org/10.1016/j.concog.2014.04.001.

Warner, H.D., Heimstra, N.W., 1971. Effects of intermittent noise on visual search tasks of varying complexity. Percept. Mot. Skills 32 (1), 219–226. https://doi.org/10.2466/pms.1971.32.1.219.

Weisberg, S., 2005. Applied Linear Regression, third ed. John Wiley & Sons, Inc., New York https://doi.org/10.1002/0471704091.

Wetter, O.E., 2013. Imaging in airport security: past, present, future, and the link to forensic and clinical radiology. Journal of Forensic Radiology and Imaging 1, 152–160. https://doi.org/10.1016/J.JOFRI.2013.07.002.

Wickelgren, W.A., 1977. Speed-accuracy tradeoff and information processing dynamics. Acta Psychol. 41 (1), 67–85. https://doi.org/10.1016/0001-6918(77)90012-9.

Wolfe, J.M., 1998. What can 1 million trials tell us about visual search? Psychol. Sci. 9 (1), 33–39. https://doi.org/10.1111/1467-9280.00006.

Wolfe, J.M., Horowitz, T.S., Kenner, N.M., 2005. Rare items often missed in visual searches. Nature 435, 439–440. https://doi.org/10.1038/435439a.

Wolfe, J.M., Horowitz, T.S., Wert, M.J. Van, Kenner, N.M., Place, S.S., Kibbi, N., 2007. Low target prevalence is a stubborn source of errors in visual search tasks. J. Exp. Psychol. Gen. 136 (4), 623–638. https://doi.org/10.1037/0096-3445.136.4.623.

Wolfe, J.M., Van Wert, M.J., 2010. Varying target prevalence reveals two dissociable decision criteria in visual search. Curr. Biol. 20 (2), 121–124. https://doi.org/10.1016/j.cub.2009.11.066.

Yu, R., Yang, L., Liu, C., 2015. Effects of target prevalence and speech intelligibility on visual search performance. Meas. Control 48 (3), 87–91. https://doi.org/10.1177/0020294015569265.