Applied Ergonomics 79 (2019) 66–75

Contents lists available at ScienceDirect

Applied Ergonomics

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

Work design for airport security officers: Effects of rest break schedules and adaptable automation

Alain Chavaillaz[a,*], Adrian Schwaninger[b], Stefan Michel[b], Juergen Sauer[a]

[a] Department of Psychology, University of Fribourg, Fribourg, Switzerland
[b] School of Applied Psychology, University of Applied Sciences and Arts Northwestern Switzerland, Olten, Switzerland

ARTICLE INFO

Keywords:
Adaptable automation
Airport security
Visual inspection
System reliability
Performance

ABSTRACT

This study investigated whether there is empirical support for the current EU regulation mandating breaks of at least 10 min after each period of 20 min continuously reviewing X-ray images in airport security screening. As a second goal, it examined whether providing more autonomy to airport security officers (in the form of spontaneous rest breaks and adaptable automation) would improve their performance and subjective state. Seventy-two student participants had to indicate the presence (or absence) of a threat item (either a gun or a knife) in a series of grey-scaled X-ray images of cabin baggage. Three work-rest schedules were examined: spontaneous breaks (i.e. participants could take breaks at any time), two 5-min breaks and two 10-min breaks during a 1-h testing session. Furthermore, half of the participants were assisted in their task by an adaptable support system offering three levels of automation: (1) no support, (2) cues indicating the presence of a potential threat item, and (3) cues indicating the exact location of a potential threat item. Results showed no performance differences between break regimes, which suggests that there may be viable alternatives to the current EU regulations. It also emerged that providing participants with adaptable automation did not lead to better detection performance but resulted in a less positive response bias than participants without automatic support. Implications for current aviation security regulations are discussed.

1. Introduction

Inspecting X-ray images of cabin baggage is a visual search and decision-making task (Koller et al., 2009; Wales et al., 2009), requiring sustained vigilance for a prolonged period of time (Warm et al., 2008). Maintaining vigilance under these conditions is difficult. For example, laboratory studies using simple target detection tasks showed performance decrements after 15 min (Teichner, 1974), but for more demanding detection tasks, such a decline can even appear after 5 min (e.g. Helton et al., 1999). Concerns about vigilance decrements are also reflected in workplace regulations. For example, the European regulation specifies that airport security officers (or screeners) shall normally not spend more than 20 min continuously reviewing X-ray images of passenger bags and they shall interrupt this activity for at least 10 min after a 20-min period (European Commission, 2015). In practice, this means that after 20 min of X-ray screening either screeners have to take a break of at least 10 min (representing the typical arrangement when working in an office-like environment conducting remote cabin or hold baggage screening) or they have to change to a different task (representing the typical arrangement when working at an airport security checkpoint). This regulation imposes a rather rigid rest break regime with limited autonomy for screeners, which represents considerable disadvantages for day-to-day operations. To our knowledge, no previous study has aimed to evaluate different work-rest schedules (including spontaneous breaks) in the domain of baggage screening at airports. Additionally, we investigated the possible benefits of giving support to the operator in the form of adaptable automation. Both, flexible work-rest schedule (or break regimes) and adaptable automation may be considered elements of operator autonomy, which represents an important concept in theories of work design (e.g. Hackman and Lawler, 1971; Hackman and Oldham, 1976; Karasek and Theorell, 1990). This question is of particular relevance given that for screeners autonomy is rather low, which is due to rather strict operating procedures at airport security checkpoints and remote screening locations (Bassetti, 2017; Kirschenbaum, 2015).

1.1. Work-rest schedules

The scheduling of rest breaks is a research question that has been examined in work psychology for about a century (e.g. Jones, 1919;

* Corresponding author. Department of Psychology, University of Fribourg, Rue de Faucigny 2, 1700, Fribourg, Switzerland.
E-mail address: alain.chavaillaz@unifr.ch (A. Chavaillaz).

https://doi.org/10.1016/j.apergo.2019.04.004
Received 17 October 2017; Received in revised form 27 December 2018; Accepted 8 April 2019
Available online 01 May 2019
0003-6870/

cited in Tucker, 2003). There are multiple benefits of rest breaks. Rest breaks can efficiently prevent performance decline (e.g. Balci and Aghazadeh, 2003; Colquhoun, 1959; Galinsky et al., 2000; Kopardekar and Mital, 1994; Steinborn and Huestegge, 2016), reduce fatigue and discomfort (e.g. Galinsky et al., 2000; Henning et al., 1997) as well as lower work-related stress hormones (Engelmann et al., 2011). They can even lead to a temporary increase in performance (e.g. Dababneh et al., 2001; Lim and Kwok, 2016; Steinborn and Huestegge, 2016). While there is little debate about the important function of rest breaks in the design of work schedules, it is more difficult to answer the question of how exactly breaks should be designed. This includes the duration and frequency of breaks (Tucker, 2003) but also the activity during the break (e.g. Steinborn and Huestegge, 2016).

Kroemer and Grandjean (2000) made a distinction between four types of breaks: scheduled breaks, hidden breaks, work-related interruptions and spontaneous breaks. In scheduled breaks, the timing and length of the work interruptions are predefined by the management and have to be respected. During hidden breaks, the employee engages in an activity that allows recovering from the strain of her/his main task (e.g. working at the metal detector of the security checkpoint instead of inspecting X-ray images). Work-related interruptions correspond to involuntary breaks due to sub-optimal work organisation or technical problems (e.g. stopping the conveyor belt of the X-ray machine when the bypass is fully loaded or reached a predefined limit of baggage). Finally, spontaneous breaks (or self-paced breaks) are work interruptions that are initiated by the employee her-/himself. Of particular relevance in the present study are scheduled and spontaneous breaks.

When examining the effectiveness of scheduled breaks, several studies showed the advantages of so-called microbreaks. These are defined as short rest periods lasting up to 10 min (Sluiter et al., 2000). The optimal break length is difficult to determine (Strongman and Burt, 2000; Tucker, 2003) and seems to depend on work characteristics (Strongman and Burt, 2000). For physical work, less frequent but longer microbreaks are more suitable. For example, workers at a meat processing plant reported less discomfort in the lower extremities after a 9-min break (every 51 min) than after a 3-min break (every 27 min, Dababneh et al., 2001). In contrast, for cognitively demanding tasks better performance and less physical discomfort were observed for shorter microbreaks. For example, Boucsein and Thum (1997) found that patent examiners were more effective when they took a 7.5-min break every 50 min than a 15-min break every 100 min. Balci and Aghazadeh (2003) observed that very short microbreaks frequently taken (e.g. 30-second break every 15 min) improved performance in data entry and mental arithmetic tasks. Surgeons benefited from additional short microbreaks (1.5–2-min break every 20–40 min) during long operations (Hallbeck et al., 2017). Overall, the research literature suggests benefits for more frequent microbreaks in cognitive tasks. However, it may be less important how long exactly a microbreak is. For example, Bennett (2015) reported no difference for students taking part in a computerised change detection task whether they had a 1-, 5-, or 9-min microbreak between two 10-min work periods.

While there is a reasonable amount of research on the effectiveness of scheduled breaks, there is much less work on spontaneous breaks. The latter may be able to compensate possible drawbacks associated with scheduled breaks (e.g., length of break may not be suitable for an optimal recovery or a scheduled break may be perceived as unnecessary because it may disrupt workflow). Henning et al. (1989) tested experienced data entry operators during six periods of 40 min, allowing them to take a self-paced microbreak in the middle of each period. The results showed that too long microbreaks were associated with higher subjective levels of fatigue and boredom. Furthermore, objective performance was systematically lower after the break, which suggests that operators may have finished their break before full recovery had been achieved. Other work showed positive effects for operators at video display terminals on the subjective level (i.e. spontaneous breaks reduced discomfort) without displaying negative effects on performance, compared to scheduled breaks (Henning et al., 1994). Similarly, microbreaks at 20-min intervals led to a reduction in operator discomfort for operators at video display terminals (McLean et al., 2001). Furthermore, worker productivity was not impaired by these additional microbreaks. Other work showed positive effects of short breaks for employees' daily work engagement (Kühnel et al., 2016). Considering the findings of the literature, spontaneous breaks could be a viable alternative to the current regulation on work-rest schedule of X-ray screening for two reasons. Spontaneous breaks seem not to affect performance and provide screeners with more autonomy in managing their work schedule.

1.2. Automation design

Increasing autonomy for baggage screeners may also be achieved through modern approaches to automation design such as adaptable automation (Scerbo, 2006). Compared to conventional baggage screening without system cues, adaptable automation allows operators to delegate tasks to the machine such as data acquisition and analysis (see Parasuraman et al., 2000). This means that such automatic systems provide a visual cue pointing at the location of the potential prohibited item (e.g. Singh and Singh, 2003). An increasing number of X-ray scanners is equipped with such automatic detection systems for explosives (e.g. Hättenschwiler et al., 2018; Wells and Bradley, 2012) and guns (e.g. Mery et al., 2013). In contrast to static automation (i.e. the kind of cuing is fixed to one automation level), adaptable automation allows operators to choose between different levels of support such as indirect cues (system indicates the presence of a prohibited item) or direct cues (system points at the exact location of a prohibited item; e.g. Chavaillaz et al., 2018). For example, operators may change the level of automation (LOA) because of increasing fatigue or select a LOA that fits their abilities. Experimental research demonstrated that there are advantages of adaptable automation over static automation (Inagaki, 2003; Parasuraman et al., 1993; Sauer et al., 2012). Implementing adaptable automation in X-ray scanners would give screeners more autonomy because they can select their preferred LOA.

Two central concepts of the present study (i.e. spontaneous breaks and adaptable automation) share a common feature. Both represent forms of work design that are characterised by an increased level of operator autonomy. Operator autonomy (and related concepts such as operator control and decision latitude) represents a core concept in the literature on work psychology (e.g. Hackman and Lawler, 1971; Hackman and Oldham, 1976; Karasek and Theorell, 1990). While spontaneous breaks refer to the temporal dimension of autonomy, adaptable automation refers to autonomy in the distribution of tasks in Gulowsen, (1972) framework. Overall, providing more autonomy seems to be beneficial to operators.

1.3. Present study

The main goal of the present study was to examine whether a break regime currently in operation at many European airports can be empirical supported. These EU regulations mandate breaks of at least 10 min after each period of 20 min continuously reviewing X-ray images (European Commission, 2015). In addition, the study aimed to examine if providing more autonomy to airport security officers in the form of spontaneous rest breaks and adaptable automation would have an impact on detection performance and the subjective operator state.

Student participants had to inspect X-ray images of passenger bags and decide whether a threat item (i.e. a gun or a knife) was present or not. Overall, all participants had to inspect X-ray images for 60 min (break time not being included). Three different break regimes were used. In the 5-min break condition, the 1-h session was split into three work periods of 20 min including compulsory two 5-min breaks between them. In the 10-min break condition, three 20-min periods were separated by two 10-min breaks. This condition is compliant with the current EU regulation (see above). Finally, in the spontaneous break condition, participants could take as many breaks as they wished but the accumulated break time was limited to 20 min. Half of the participants worked without automatic support, whereas the other half had at their disposal a support system offering the possibility to choose between three levels of support (no cue, indirect cue, direct cue) and could change levels at any time. Several outcome variables were measured, such as detection performance, response bias, response time, perceived workload, self-confidence in task completion, and trust in automation. Prior to the main testing session, participants' inherent detection performance was measured to control for differences in aptitude across experimental groups.

With regard to the effectiveness of break type, the research literature does not provide evidence for the superiority of the break regime mandated by the current EU regulations. Similarly, there is no unequivocal evidence in the literature that spontaneous breaks lead to better performance. Therefore, we did not make a specific prediction about the effect of breaks but rather examined the question in an exploratory way. If it emerged that there was no difference between break conditions, this would count in favour of a more liberal break regime, suggesting that all break types might represent acceptable options. With regard to the effects of automation, we expected that detection performance was better with automatic support than without due to the 'second opinion' offered by the automatic system, resulting in better decisions of the screener. We also expected that the experimental condition 'spontaneous break and adaptable automation' would result in a more positive subjective reaction (e.g., perceived workload) than other conditions because it best implemented the recommendations of work design theory with regard to operator autonomy.

2. Method

2.1. Participants

Seventy-two students (62 females and 10 males) from the University of Fribourg, aged from 19 to 44 years ($M = 22.39$, $SD = 4.14$), participated in this experiment. They received either course credits or CHF 30 as compensation in return for their participation. The number of participants was determined based on an a priori power analysis (alpha = .05 and beta = 0.85 with a medium to large effect size expected).

2.2. Simulation

A modified version of the Luggage Inspection Simulation (LIS), developed by Chavaillaz et al. (2018), was used in the current study. The simulation allowed modelling an imperfect support system (about 80% reliability) with three levels of automatic support (LOA), based on the work of Goh et al. (2005). At LOA 1, participants completed their task without automation support (see Fig. 1A). At LOA 2, the system provided an indirect cue about the presence or absence of a potential target item. When automation suggested the presence of a target, the whole piece of baggage was surrounded by a rectangular red frame (see Fig. 1B). At LOA 3, the system provided a direct cue, namely the potential threat item was framed by a red rectangle (see Fig. 1C). At both LOA 2 and 3, the word "Target" appeared in the top part of the X-ray image. When automation declared a bag to be harmless, the words "No Target" appeared in green. Automation failures were of two types: misses (for target-present trials) and false alarms (for target-absent trials). No miscues were used (i.e. an object other than the target is cued in a target-present trial). Participants were informed that automation might fail but neither about its exact reliability nor the nature of the failures.

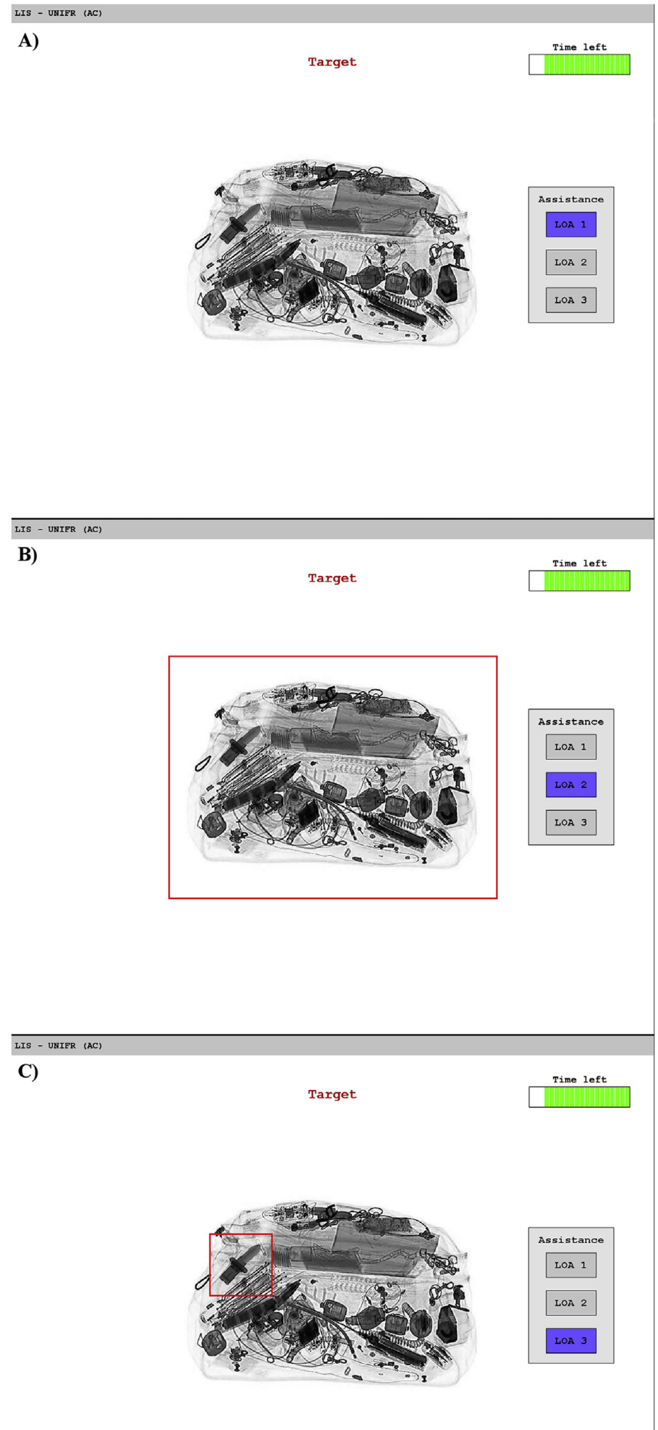

Fig. 1. Interface of LIS (Luggage Inspection Simulation) in adaptable automation mode. In the top right corner, the countdown timer is displayed, and on the right to the X-ray image, the control panel for selecting levels of automation (LOA) is shown. Panel A: At LOA 1, no automation support is provided. Panel B: At LOA 2, the piece of baggage is surrounded by a red frame and the word "Target" is displayed when automation has detected the presence of a threat item. Panel C: At LOA 3, the threat item is surrounded by the red frame and the word "Target" is displayed. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.3. Apparatus and stimuli

The Luggage Inspection Simulation (LIS) was run by a Matlab script using the Psychtoolbox (Brainard, 1997; Kleiner et al., 2007; Pelli,

1997) on a Dell PC with Microsoft Windows 7 operating system. Participants were seated at an approximate distance of 0.60 m in front of a 17” LCD flat screen (resolution of 1280 × 1024 pixels and refresh rate of 60 Hz). They were free to move their head. The height and width of the X-ray images of cabin bags were between 12° and 14° of visual angle.

The pretest consisted of 64 target-present and 64 target-absent X-ray images of cabin baggage taken from the X-Ray Object Recognition Test, which is used for selecting screeners for jobs in airport security (X-Ray ORT; Hardmeier et al., 2005; Schwaninger et al., 2005). A pool of 831 real X-ray images of cabin baggage was used in the main test. 128 of these bags contained a threat item whereas the other bags were harmless. This ratio was based on previous baggage screening studies (e.g. Goh et al., 2005; McCarley et al., 2004).

For both testing phases (pretest and main test), threat items were virtually merged into X-ray images of passenger bags by experts (i.e. former screeners) using empirically validated software algorithms (Bastian et al., 2009). Knives and guns were used as targets since they are more easily recognisable by student participants than other threat items (e.g. explosives or electric shock devices). Since student participants do not know the meaning of colour in X-ray images, grayscale X-ray images were used.

2.4. Design

A 3 × 2 × 3 experimental design was used in the present study (see Fig. 2). The first between-subjects factor with three levels was the break condition. Each participant had to inspect X-ray images of cabin baggage for 60 min (break time not included). In the spontaneous break condition, participants had a total of 20 min of break time available, which they could take whenever they wished (see Procedure for more details). In the 5-min break condition (5-min), participants worked for three periods of 20 min interrupted by two breaks of 5 min each, whereas in the 10-min break condition (10-min), the two breaks lasted 10 min each. The second between-subjects factor, automatic support, had two levels. In the adaptable automation condition, participants were provided with a support system that allowed them to change the LOA at any time. In the no-automation condition, participants had to carry out their task without automatic support. Time-on-task was used as a third within-subjects factor for the analysis, with three levels (i.e. three blocks of 20 min). Participants were randomly assigned to one of the six experimental groups (i.e. 12 participants per group).

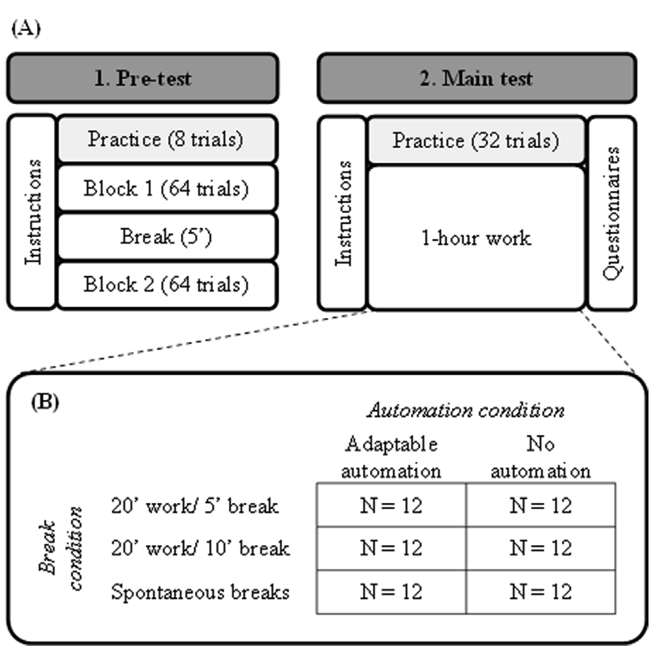

Fig. 2. Illustration of the procedure (A) and the experimental design (B).

2.5. Dependent measures

Three measures were used to assess participant performance. First, participant ability to indicate the presence or absence of a threat item (referred to as *detection performance*) was measured by d’ = z (hit rate)-z (false alarm rate). Second, the *response bias* c = −0.5*[z (hit rate) + z (false alarm rate)] was used to assess participant tendency to answer positively or negatively. For example, a positive response bias means that a screener, if unsure, would tend to report the presence of a target (i.e. risking a false alarm) rather than the absence of a target (i.e. risking a miss). In both formulae, z refers to the inverse of the cumulative distribution function of the standard normal distribution (Green and Swets, 1966). Third, for each image *response time* was measured in milliseconds (but reported in seconds) by taking the time from image onset until a response was made.

For participants in the adaptable automation condition, automation management was determined by the mean automation level chosen during a trial (*average LOA*) and the propensity to switch between LOAs (*LOA stability*). Furthermore, the propensity to acknowledge automation recommendation was assessed by two measures introduced by Meyer (2001): *compliance* and *reliance*. We used the same method for measuring the two variables as Rice and McCarley (2011) did. Compliance refers to the participant's tendency to follow automation recommendation when it indicates the presence of a target. Reliance corresponds to the participant's tendency to follow automation recommendation when it indicates the absence of a target.

Trust was assessed by the 12-item questionnaire “Checklist for Trust between People and Automation” (CTPA) developed by Jian et al. (2000). Participants had to rate each item on a 7-point Likert scale. An item example was ‘The system is reliable’.

A purpose-built item was used for measuring *self-confidence*, which was adapted from Lee and Moray (1992). Participants had to answer the following question on a 10-point Likert scale: ‘How confident were you in your ability to detect dangerous objects?’.

The 6-item questionnaire NASA-TLX (Hart and Staveland, 1988) was used to assess *workload*. Each of the six items measuring a specific workload aspect (i.e. mental demands, physical demands, temporal demand, performance, frustration, and effort) was rated on a 20-point Likert scale. For instance, the item assessing performance was: ‘How successful were you in accomplishing what you were asked to do?’

2.6. Procedure

The experiment consisted of two distinct phases (pretest and main test). The pretest aimed to control for possible differences between participants’ inherent ability and aptitude to detect visual stimuli across the experimental groups. The main test aimed to assess the influence of break regime and automation on outcome variables. A summary of the procedure and experimental design is depicted in Fig. 2.

Participants had to decide for each X-ray image of cabin baggage whether it contained a threat item or not (yes-no task in signal detection theory; Macmillan and Creelman, 2005). They were instructed to respond as accurately and quickly as possible by clicking on the left or right mouse button. The mapping between the mouse button and the type of response (target absent vs. target present) was counterbalanced across participants.

2.6.1. Pretest

The pretest started with instructions describing the task and the response modalities. Participants then completed a short practice block of eight trials (including feedback about their performance, target type and target location), followed by two experimental blocks of 64 trials (without feedback) separated by a 5-min break. The two sets of

Table 1
F-value, significance level and effect size for the main effects of break and automatic support conditions and their interaction.

Variable	Break			Automation			Break X Automation		
	F^a	p	$\eta^2_{partial}$	F^b	p	$\eta^2_{partial}$	F^c	p	$\eta^2_{partial}$
Performance									
Detection ability	0.025	.975	.001	0.003	.958	< .001	1.614	.207	.047
Response bias	0.874	.422	.026	**18.305**	**< .001**	**.331**	1.256	.291	.037
Response time									
Target present	1.472	.237	.043	2.281	.136	.033	1.368	.262	.040
Target absent	3.089	.052	.086	1.194	.279	.018	2.505	.089	.071
Efficiency	1.758	.180	.051	3.577	.063	.051	2.742	.072	.077
Subjective measures									
Trust	0.796	.460	.046	–	–	–	–	–	–
Perceived workload	1.476	.236	.043	**6.300**	**.015**	**.087**	5.127	.009	.134
Self-confidence	0.048	.954	.001	2.331	.132	.034	**3.631**	**.032**	**.099**
Use of automation									
Average LOA	0.139	.871	.008	–	–	–	–	–	–
LOA switch frequency	1.836	.175	.100	–	–	–	–	–	–
Reliance	1.061	.358	.060	–	–	–	–	–	–
Compliance	0.821	.449	.047	–	–	–	–	–	–

Note. Significant effects are in boldface. LOA = Level of automation; [a/c] df: 2, 66; [b] df: 1, 66.

potential targets (guns and knives) were presented for 10s each at the beginning of the practice block and before the first experimental block to make participants familiar with the target items.

Each trial started with a fixation cross appearing on screen for 500 ms. The X-ray image was presented for 4s and participants had up to 20s to respond. If no response was given within 20s, the trial stopped being counted as a target-absent response. Participants were informed about the remaining time by a countdown timer in which the number of vertical bars indicated the number of seconds left to respond (see Fig. 1). A blank screen appeared for 500 ms when the X-ray image disappeared.

2.6.2. Main test

At the beginning of the main test, participants received written and oral instructions about the experimental condition to which they were assigned. Participants in the automation condition practised with the support system during the training block consisting of 32 trials (50% target-present trials). During the experimental blocks, participants inspected 614 X-ray images on average (SD = 126.52) including all 128 target-present images (about 20% target-present trials). When the experimental task was completed, participants filled in several questionnaires (i.e. trust towards automation, perceived automation reliability, self-confidence in their ability to achieve the task, and subjective workload). In total, participants needed 70 min to complete the main test in the 5-min condition, 80 min in the 10-min condition and on average 74 min (SD = 6) in the spontaneous break condition.

The trial sequence was identical to the pretest with two exceptions. First, the X-ray image remained on screen for up to 20s or until a response was given. Second, after the X-ray image had disappeared, a blank screen appeared for 1500 ms in the 5-min and 10-min conditions, whereas, in the spontaneous break condition, the button 'Take a break' was available in the middle of the screen for 1500 ms. When pressed, a screen appeared displaying a digital countdown indicating the remaining break time and a button 'Continue', which, when pressed, allowed participants to resume the test.

2.7. Data analysis

To examine whether participants' detection performance (d') was equivalent between experimental groups, data from the pretest were analysed. A one-way ANOVA including all groups was carried out on detection performance in the pretest. No significant difference was observed at an alpha level of 0.20 (null hypothesis testing), F (5,66) = 0.410, p = .840, $\eta^2_{partial}$ = .030. The Levene's test of variance homogeneity was not significant, F (5,66) = 1.063, p = .389. These results confirmed that the inherent detection abilities of participants were not significantly different in the six experimental groups.

The data set of each participant was checked for possible outliers in response time (separately for target-present and target-absent trials). All data points that exceeded three standard deviations were removed.

Different ANOVA models were used in the data analysis due to differences in taking measurements for the different variables. For each performance measure, a three-way ANOVA was carried out to assess the influence of break regime, automation level, and time-on-task. For self-confidence and perceived workload, a two-way ANOVA was carried out to evaluate the impact of break regime and automation level since these dependent variables were not measured as a function of time. For the measures referring to 'use of automation', the influence of break regime and time-on-task was assessed by a two-way ANOVA (because the dependent variable could only be measured in the adaptable automation condition). For trust, a one-way ANOVA was carried out with break regime as the only factor (this measure was only collected once per session and could only be sensibly applied in the condition 'adaptable automation'). Bonferroni corrections were used in case of multiple post-hoc comparisons and the correction proposed by Dunlap et al. (1996) for Cohen's d was applied for dependent sample t-tests. In order to improve readability of the results section, the outcomes of most statistical tests will be reported in tables. As in Ostarek et al. (2019), the data were also analysed with the Bayesian approach (using JASP; JASP Team, 2018) by comparing the fit of the data with the null model and the alternative models (including main effects and interactions). Only the model with the largest evidence (BF_{10}) is reported. When a discrepancy between the frequentist and the Bayesian approach occurred, the simplest model was favoured following the rule of parsimony.

3. Results

3.1. Performance

Detection performance. Participants showed an average detection performance (d') of 2.08 (SD = 0.40). The ANOVA revealed no effect of break condition, none of automatic support, and no significant interaction between the two factors was observed (see Table 1). The data analysis also showed improvements in detection performance from block 1 (M = 1.91, SD = 0.37) to block 2 (M = 2.14, SD = 0.42), which then stabilised in block 3 (M = 2.18, SD = 0.55). This main effect of time-on-task was significant (see Table 2), with pairwise comparisons confirming that only the increase from block 1 to block 2 was

Table 2
F-value, significance level and effect size for the main effect of time-on-task and its interaction with break and automatic support conditions.

Variable	ToT			ToT X Break			ToT X Automatic support			3-Way Interaction		
	F	*p*	$\eta^2_{partial}$	*F*	*p*	$\eta^2_{partial}$	*F*	*p*	$\eta^2_{partial}$	*F*	*p*	$\eta^2_{partial}$
Performance												
Detection performance	**20.664**[a]	**< .001**	**.238**	2.260[b]	.073	.064	3.065[a]	.055	.044	1.315[b]	.270	.038
Response bias	**32.582**[c]	**< .001**	**.331**	1.854[d]	.134	.053	0.572[c]	.566	.009	0.091[d]	.975	.003
Response time												
Target present	**26.292**[e]	**< .001**	**.285**	2.257[f]	.066	.064	0.441[e]	.644	.007	**3.242**[f]	**.014**	**.089**
Target absent	**89.376**[g]	**< .001**	**.575**	**3.108**[h]	**.034**	**.086**	0.198[g]	.739	.003	2.250[h]	.092	.064
Use of automation												
Average LOA	0.225[i]	.799	.007	0.623[j]	.617	.036	–	–	–	–	–	–
LOA switch frequency	**4.107**[k]	**.036**	**.111**	2.462[l]	.054	.130	–	–	–	–	–	–
Reliance	**19.235**[m]	**< .001**	**.368**	1.185[n]	.326	.067	–	–	–	–	–	–
Compliance	**16.038**[o]	**< .001**	**.327**	0.479[p]	.744	.028	–	–	–	–	–	–

Note. Significant effects are in boldface. ToT = Time-on-task; LOA = Level Of Automation*;* [a] df = 1.995, 131.307; [b] df = 3.991, 131.307; [c] df = 1.865, 123.106; [d] df = 3.730, 123.106; [e] df = 2132; [f] df = 4, 132; [g] df = 1.521, 100.381; [h] df = 3.042, 100.381; [i] df = 1.822, 60.131; [j] df = 3.644, 60.131; [k] df = 1.505, 49.677; [l] df = 3.011, 49.677; [m] df = 1.852, 61.100; [n] df = 3.703, 61.100; [o] df = 2, 66;[p] df = 4, 66.

significant, $t_{B1\text{-}B2}$ (71) = −5.692, p < .001, D = 0.949. None of the interactions was found to be significant (see Table 2). The model with only the main effect of time-on-task achieved the highest evidence in the Bayesian ANOVA ($BF_{10} = 2.942 \times 10^5$).

Response bias. Regarding the response bias, there was no main effect of break. However, participants without automatic support responded significantly more frequently that the target was present (M = 0.58, SD = 0.31) than participant supported by automation (M = 0.28, SD = 0.31, see Table 1). This result indicated a more positive response bias for the no-automation group. The interaction between break and automatic support was not significant (see Table 1). Furthermore, the main effect of time-on-task was significant (see Table 1 and Fig. 3). Further analyses showed that with increasing time-on-task participants responded more often that a target was present, indicated by three significant pairwise comparisons, all t (71) > 4.000, p < .001, D > 0.667. All other interactions were not significant (see Table 2). This model with the main effects of automation and time-on-task achieved the largest evidence in the Bayesian ANOVA ($BF_{10} = 1.113 \times 10^{12}$).

Response time. As expected, participants were significantly faster to indicate the presence (M = 2.22, SD = 0.75) than the absence of a target in the X-ray image (M = 3.43, SD = 1.60), F (1,66) = 96.751, p < .001, $\eta^2_{partial}$ = .594. For this reason, the response times for target-present and target-absent trials were analysed separately.

Response time: target present. The response times for target-present trials are shown in Fig. 4. The three-way ANOVA revealed no effect of break conditions, none of automatic support, and no significant interaction between the two factors (see Table 1). However, it showed a significant main effect of time-on-task (see Table 2). Pairwise comparison revealed decreasing response times from block 1 (M = 2.45s,

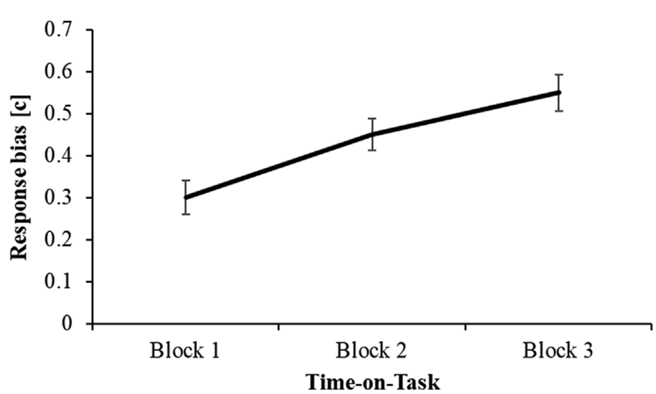

Fig. 3. Mean response bias (and standard errors) as a function of time-on-task.

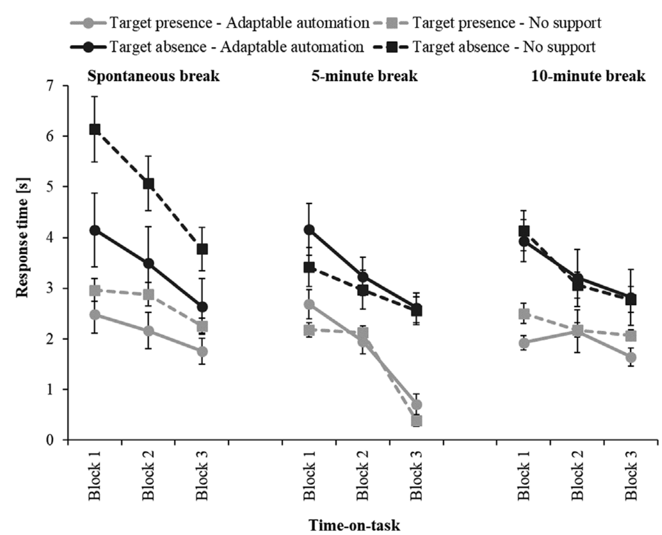

Fig. 4. Mean response time (and standard errors) for target-present and target-absent trials as a function of break condition and automatic support.

SD = 0.87) to block 2 (M = 2.24s, SD = 0.98) and block 3 (M = 1.93s, SD = 0.63), with all pairwise comparisons being significant, all t (71) > 2.828, p > .018, D > 0.247. Furthermore, the three-way interaction was significant. Participants in spontaneous break condition showed a significant linear trend of performance improvements in response times as the testing session progressed (see Table 3), with no interaction between time-on-task and automatic support being observed (see Table 2). In the 5-min condition, there was also a significant main effect of time-on-task (see Table 3), which took different directions for the two automatic support conditions (see Table 2 for the significant interaction). In the no-automation group, participants significantly improved their performance from block 3 to block 2, t (11) = 2.819, p = .31, D = 1.151, whereas the other pairwise comparisons showed no such effect, t's (11) > 0.05. In the automation group, participants improved their performance from block 1 to block 2 and 3 (all t's > 4.4, p < .001, D's > 1.8). In the 10-min break condition, there were no significant improvements of performance (see Table 3). All other interactions were not significant (see Table 2). In contrast to the frequentist approach, the Bayesian analysis showed the largest evidence for a simpler model with only a main effect of time-on-task ($BF_{10} = 9.867 \times 10^6$), while there was no evidence for the 3-way interaction model suggested by the frequentist approach (BF_{10} = 0.366).

Response time: target absent. Target-absent response times are

Table 3
F-value, significance level and effect size for the main effect of time-on-task and automatic support and their interaction for each break condition for target-present response time.

Break condition	ToT			Automation			ToT X Automation		
	F	*p*	$\eta^2_{partial}$	*F*	*p*	$\eta^2_{partial}$	F	p	$\eta^2_{partial}$
Spontaneous break	**20.778**	**< .001**	**.486**	2.359	.139	.097	0.655	.524	.029
5-min break	**14.318**	**< .001**	**.394**	0.232	.635	.010	**6.239**	**.013**	**.221**
10-min break	3.064	.073	.122	1.670	.210	.071	1.645	.210	.070

Note. Significant effects are in boldface. ToT = Time-on-task; [a] df = 2, 30; [b] df = 1, 30.

presented in Fig. 4. The ANOVA showed no difference between break regimes (see Table 1). There was no effect of automatic support and no interaction between break regimes and automatic support (see Table 1). Similar to target-present trials, the main effect of time-on-task was significant (see Table 2). This was because response times decreased continuously across the three blocks (block 1: $M = 4.32$ s, $SD = 1.97$; block 2: $M = 3.51$s, SD = 1.81; block 3: $M = 2.87$ s, $SD = 1.45$), with all pairwise comparisons being significant, all t (71) > 8.089, $p < .001$, D > 1.348. Furthermore, there was an interaction between break regimes and time-on-task (see Table 2). This was because comparisons between break conditions in each block revealed that in block 1 participants had significantly slower response times in the spontaneous break regimes than in the 5-min break condition, t (22) = 2.560, $p = .038$, D = 1.045 (all other *t's* < 2.1, p > .10). In blocks 2 and 3, this difference had disappeared, all t (22) < 2.380, $p > .061$, D < 0.972. None of the other interactions was significant (see Table 2). This model with the interaction between time-on-task and break regime showed the largest evidence in the Bayesian ANOVA ($BF_{10} = 3.910 \times 1021$). According the parsimony rule, the Bayesian model was retained.

3.2. Subjective measures

Trust. Participants working with automation had on average a trust level of $M = 3.29$ ($SD = 0.80$) on a scale from 1 to 7. There was no effect of break (see Table 1). This null model achieved the highest evidence in the Bayesian ANOVA (all other models had $BF_{10} < 1$).

Perceived workload. The data in Fig. 5 show that there was no effect of break on perceived workload (see Table 1). The data analysis revealed significantly higher ratings for participants working with automation than without (see Table 1). There was also a significant interaction between break conditions and automatic support (see Table 1). This was because only participants in the 5-min break condition had a significantly higher workload level with automatic support than without, t (22) = 3.937, $p = .020$, D = 1.607. This was in contrast to the two other break conditions, which showed no difference between automation and no-automation conditions, both t (22) < 0.906, $p > .136$, D < 0.370. This model is also supported by the Bayesian approach ($BF_{10} = 5.048$).

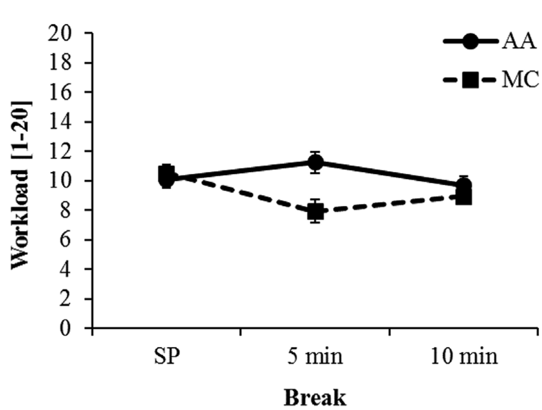

Fig. 5. Mean level of perceived workload (and standard errors) as a function of break condition and automatic support presence.

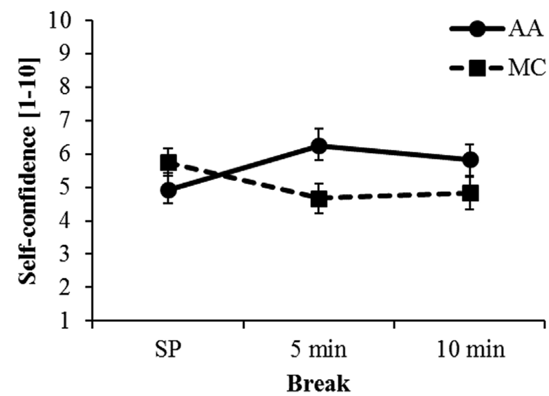

Fig. 6. Mean level of self-confidence (and standard errors) as a function of break condition and automatic support.

Self-confidence. The ratings for self-confidence are presented in Fig. 6. The two-way ANOVA revealed no main effect of break and none of automatic support (see Table 1). However, the interaction between factors was significant (see Table 1). Post-hoc analyses revealed that participants in the 5-min break condition had a significantly higher level of self-confidence under automation than without support, t (22) = 2.391, $p = .020$, D = 0.976. In the two other break conditions, no such difference was found between automation and no-automation conditions, both t (22) < 1.511, $p > .136$, D < 0.617. This effect was not confirmed by the Bayesian analysis, which favoured a null model (all other models had $BF_{10} < 1$). According to the parsimony rule, the Bayesian model was retained.

3.3. Use of automation

Average LOA. On average, participants used a mean LOA of about $M = 2.47$ ($SD = 0.68$), which showed that participants generally made use of the automatic support offered by the system. The analysis revealed no significant differences between experimental conditions (i.e. no main effect of break or time-on-task and no interaction between the two factors). The results of the frequentist statistical tests are reported in Tables 1 and 2 This null model was achieved the highest evidence in the Bayesian ANOVA (all other models had $BF_{10} < 1$).

LOA change frequency. Overall, participants changed LOA not very frequently (M = 0.06, SD = 0.11). The break regime did not show an effect on LOA changes (see Table 1). The analysis revealed a significant main effect of time-on-task (see Table 2). This was because participants changed LOA less frequently as the testing session progressed (block 1: $M = 0.08$, $SD = 0.15$; block 2: $M = 0.05$, $SD = 0.12$; block 3: $M = 0.04$, $SD = 0.11$). This change in automation management behaviour was found to be a significant linear trend, F (1,33) = 4.928, $p = .033$, $\eta^2_{partial} = .130$. Finally, there was no significant interaction (see Table 2). The Bayesian analysis showed anecdotal evidence for a more complex model including the interaction of time-on-task and

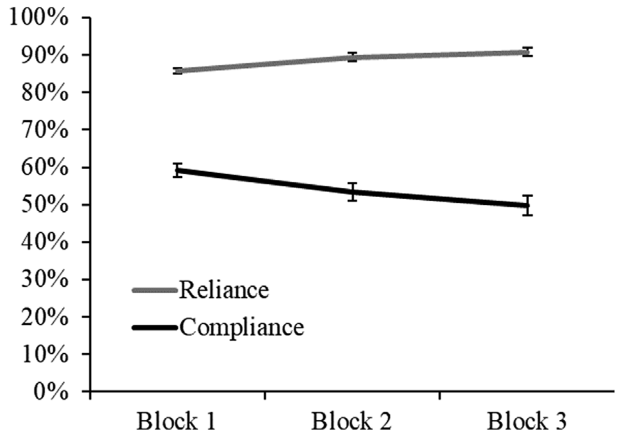

Fig. 7. Mean percentage of reliance and compliance (and standard errors) as a function of time-on-task.

break regime (BF_{10} = 1.813). The second best model including only time-on-task (as suggested by the frequentist analysis) achieved similar evidence in the Bayesian ANOVA (BF_{10} = 1.697). Since there was no evidence in favour of the more complex model when compared to the second model (BF_{10} = 1.050), the latter was retained following the parsimony rule.

Reliance. Participant reliance on automation was very high overall (see Fig. 6). The ANOVA revealed that it was not affected by break conditions (see Table 1). When automation indicated the absence of a target, participants heeded the advice more often with increasing session duration (see Fig. 6). This was confirmed by a significant main effect of time-on-task (see Table 2). Post-hoc analysis showed that towards block 1 was significantly different from the two other blocks, both $t\,(35) > 4.173, p < .001$, D > 0.984. No interaction was observed (see Table 2). The Bayesian analysis also resulted in the largest evidence for the time-on-task model (BF_{10} = 5.193 × 104).

Compliance. When automation indicated the presence of a target, participants showed similar compliance levels in each break condition (see Table 1). As shown in Fig. 7, their rate of compliance diminished over time (see Table 2). Pairwise comparisons indicated that this reduction in compliance was significant from block 1 (M = 59.24%, SD = 11.02) to block 2 and block 3, both t (35) > 3.594, p < .001, D > 0.847 but not between block 2 and 3, t (35) = 2.315, p = .081, D = 0.546. Finally, there was no significant interaction (see Table 2). As for reliance, the Bayesian analysis resulted in the largest evidence for the model including only time-on-task (BF_{10} = 1.103 × 104).

4. Discussion

The first goal of the present study was to examine whether there is empirical support for the rather strict EU regulation on work-rest schedules that are currently in place for airport security screeners. The results provided no such empirical support for the benefits of obligatory 10-min breaks compared to alternative break regimes. The second goal was to examine whether providing more autonomy to screeners would impact their performance and subjective state. Offering more autonomy to screeners (either in the form of spontaneous rest breaks or adaptable automation) showed no performance difference compared to traditional forms of work design. The results were almost identical for the classical frequentist and Bayesian approaches of statistical analysis.

Overall, the manipulation of break regime did not affect performance variables and had only few effects on subjective measures. This has two important implications. First, the present study did not find empirical evidence for the benefits of compulsory 10-min breaks over 5-min breaks (after 20 min of continuously reviewing X-ray images, as it is required by current EU regulations), with the latter being equally effective. Second, the break regime involving more operator autonomy (i.e. spontaneous breaks) was of similar effectiveness compared to the two fixed-break regimes. We believe that both findings are of relevance since they provide first evidence that alternative design options for work-rest schedules (i.e. shorter breaks and more flexible ones seem to be of equal effectiveness compared to the one currently prescribed by workplace regulations).

Providing participants with more autonomy in their work schedule by being able to make use of spontaneous breaks did not influence their performance in comparison to fixed work-rest regimes. Similar findings were found in previous studies (Henning et al., 1989, 1994). Given the benefits of increasing autonomy for operators as suggested by the job characteristic model (Hackman and Lawler, 1971; Hackman and Oldham, 1976; Karasek and Theorell, 1990), the effectiveness of using flexible breaks in baggage screening may warrant further exploration. However, we need to take into account important differences in work design for remote cabin and hold baggage screening compared to cabin baggage screening at security checkpoints. For the former, the implementation of spontaneous breaks may be easier because screeners work rather independently from other staff and they could take breaks without having much impact on baggage flow and overall work efficiency. In contrast, when cabin baggage screeners at the security checkpoint take a spontaneous break, it will have considerable repercussions for colleagues (e.g. they would have to replace them) and passengers alike (e.g. their waiting time may increase). Being aware of these disruptions, it may discourage checkpoint screeners to take a break (even if they felt they needed one), increasing the risk of an inappropriate recovery management. Therefore, it may be better to examine the possibilities of using spontaneous breaks first in the domain of remote cabin and hold baggage screening before moving on to checkpoint screening.

One minor effects of break regime were also observed. In the 5-min break condition, participants working with the adaptable automation reported to be more self-confident in their ability to detect a target than participants without support. The support of automation may have given participants the impression that they were overall more efficient than those participants working without automation. At the same time, they reported higher workload levels with automation than without when having 5-min breaks. This may be associated with the additional cognitive load involved in managing adaptable automaton (i.e. having to decide whether to change LOA or not, and which LOA to choose). However, this seems to have no impact on work quality since all experimental groups displayed similar performance levels.

Examining the effects of adaptable automation provided three important findings. (a) In both conditions (i.e. with and without automation), a positive response bias was observed (i.e. if unsure, participants reported that there was a target rather than there was none). In the context of baggage screening, it makes sense to err on the side of caution due to the more critical consequences of failing to identify a target than to generate a false alarm by opening a harmless piece of baggage. This positive response bias was significantly less pronounced when automated was provided than when it was not. Such an effect of automation on response bias was also observed in previous work (Chavaillaz et al., 2018) and may be explained by the following mechanism. Under the no-support condition, participants could only rely on their own judgement because a ‘second opinion’ from the automation was unavailable. This may have resulted in participants preferring to err on the side of caution on a higher number of instances. Conversely, when automatic support was provided, the positive response bias was reduced because of the ‘second opinion’ being offered about the presence of a potential threat item. (b) The presence of an adaptable support system (which provided operators with more autonomy) did not impair performance but it did not provide the expected benefits either. Participants supported by adaptable automation showed similar performance (in terms of detection performance and response time) as participants working without support. These observations confirm recent findings from baggage screening which did not show better detection performance for adaptable automation than unaided screening (Chavaillaz et al., 2018). It is not entirely clear why the expected benefits of adaptable automation failed to materialise. A contributing factor may have been the increased workload under adaptable automation (as indicated by participant ratings) compared to the unaided condition. There may have been additional cognitive resource requirements when participants were faced with a decision from the automation that was different from their own or when they needed to check whether it would be advantageous to change their current LOA (e.g. Bailey et al., 2006; Kaber and Riley, 1999). Interestingly, there may be domain-specific effects with regard to the effectiveness of adaptable automation. In other work domains (such as process control) advantages of adaptable automation were found (Sauer et al., 2012). It may be related to the level of uncertainty associated with the decision. In process control, it is often more difficult to decide what system fault may be at the root of an unsafe state. In the present study, this may have been much simpler because guns and knives were used as prohibited items. If ambivalent stimuli (e.g. improvised explosive devices) were used, the benefits of adaptable automation may be higher. (c) With regard to the use of adaptable automation, participants preferred to work with automatic support rather than without it, as indicated by an average LOA of about two and a half (out of a maximum of score of three). These results confirm findings from a recent study on baggage screening (Chavaillaz et al., 2018). Interestingly, even though the cues in LOA 3 provided the exact location of the potential threat item, some participants preferred to work with the less powerful LOA 2. Furthermore, the overall prevalence of LOA switches indicates that participants relied on the same LOA for most of the experimental session. These findings are consistent with results obtained in process control (Sauer and Chavaillaz, 2017). They provide further evidence that adaptable automation is primarily used to accommodate for participant's individual preferences rather than for within-session adaptations to accommodate for changes in workload.

Some time-on-task effects on performance were observed. During the later parts of the experiment, participants committed fewer errors in the detection of threat items and took their decisions faster. Consequently, the amount of baggage inspected increased across blocks. This is consistent with previous findings (e.g. Colquhoun and Edwards, 1970; McCarley et al., 2004; Salthouse and Somberg, 1982). It reflects the typical practice effect in experimental research using non-expert participants since continuous learning is taking place before a stable baseline is reached. More interesting may be that during the course of the experiment, there was an increase in reliance (i.e. the tendency to follow automation recommendations when it had indicated that there was no target) and a decrease in compliance (i.e. the tendency to follow automation recommendations when it had indicated that there was a target). High levels of compliance and reliance that exceed system reliability are usually problematic since they may increase the number of false alarms (for compliance) and misses (for reliance). These diverging trends represent a finding that defies any straightforward explanation. It may be related to the higher prevalence of target-absent images (about 80%) compared to target-present ones (about 20%). Participants gained more experience with the former than the latter, which may have changed the level of trust they placed into the system.

This study has a number of limitations. First, in our experiment, only 1 h of inspecting X-ray images of passenger bags was examined. A full working day needs to be modelled to examine the potential benefits of alternative work-rest schedules over a longer time. There are hints that spontaneous breaks may be more beneficial towards the end of a working day than at the beginning. For example, a study showed that self-initiated short breaks in the afternoon boosted daily work engagement whereas taking them in the morning failed to do so (Kühnel et al., 2016). Second, our study was conducted with novices in the form of student participants and we do not know whether the same effects would have also emerged for experienced screeners. For example, screeners are expected to have higher motivation to find threat items than student participants, which may affect outcome measures (Clark et al., 2012). On the other hand, in the current climate of enhanced terrorist threat, the relevance of such research is obvious and may have contributed to increased participant motivation, leading to only a marginal difference between the two groups.

Despite these limitations, the present study provides important insights since it was the first study that investigated work-rest schedules and adaptable automation in visual inspection of cabin baggage. It also carries implications for current EU aviation security regulations because less strict break regimes were found to be of equal effectiveness. Clearly, before changes to the EU regulations on rest breaks for screeners can be taken into consideration, this issue needs to be examined in further empirical studies, which should notably involve the testing of real airport security officers.

Acknowledgements

This study was supported by the Swiss National Science Foundation (Grant no: 100014_149184). Many thanks are also due to Naomi Leitenberg and Guiseppina El Harti for their help with the research project.

References

Bailey, N.R., Scerbo, M.W., Freeman, F.G., Mikulka, P.J., Scott, L.A., 2006. Comparison of a brain-based adaptive system and a manual adaptable system for invoking automation. Hum. Factors 48 (4), 693–709. https://doi.org/10.1518/001872006779166280.

Balci, R., Aghazadeh, F., 2003. The effect of work-rest schedules and type of task on the discomfort and performance of VDT users. Ergon 46 (5), 455–465. https://doi.org/10.1080/0014013021000047557.

Bassetti, C., 2017. Airport security contradictions: interorganizational entanglements and changing work practices. Ethnography 1 (4). https://doi.org/10.1177/1466138117696513. 146613811769651.

Bastian, C.C. von, Schwaninger, A., Michel, S., 2009. The impact of color composition on X-ray image interpretation in aviation security screening. In: Proceedings. 2009 International Carnahan Conference on Security Technology (ICCST), Zurich, Switzerland. IEEE Service Center, Piscataway, NJ, pp. 201–205 5/10/2009 - 8/10/2009.

Bennett, A., 2015. Take Five? Examining the Impact of Microbreak Duration, Activities, and Appraisals on Human Energy and Performance, Richmond, VA.

Boucsein, W., Thum, M., 1997. Design of work/rest schedules for computer work based on psychophysiological recovery measures. Int. J. Ind. Ergon. 20 (1), 51–57. https://doi.org/10.1016/S0169-8141(96)00031-5.

Brainard, D.H., 1997. The psychophysics toolbox. Spatial Vis. 10, 433–436.

Chavaillaz, A., Schwaninger, A., Michel, S., Sauer, J., 2018. Automation in visual inspection tasks: X-ray luggage screening supported by a system of direct, indirect or adaptable cueing with low and high system reliability. Ergon 1–35. https://doi.org/10.1080/00140139.2018.1481231.

Clark, K., Cain, M.S., Adamo, S.H., Mitroff, S.R., 2012. Overcoming hurdles in translating visual search research between the lab and the field. In: The Influence of Attention, Learning, and Motivation on Visual Search. Springer, New York, NY, pp. 147–181.

Colquhoun, W.P., 1959. The effect of a short rest-pause on inspection. Ergon 2 (4), 367–372. https://doi.org/10.1080/00140135908930451.

Colquhoun, W.P., Edwards, R.S., 1970. Practice effects on a visual vigilance task with and without search. Hum. Factors 12 (6), 537–545. https://doi.org/10.1177/001872087001200603.

Dababneh, A.J., Swanson, N., Shell, R.L., 2001. Impact of added rest breaks on the productivity and well-being of workers. Ergon 44 (2), 164–174. https://doi.org/10.1080/00140130121538.

Dunlap, W.P., Cortina, J.M., Vaslow, J.B., Burke, M.J., 1996. Meta-analysis of experiments with matched groups or repeated measures designs. Psychol. Methods 1 (2), 170–177. https://doi.org/10.1037/1082-989X.1.2.170.

Engelmann, C., Schneider, M., Kirschbaum, C., Grote, G., Dingemann, J., Schoof, S., Ure, B.M., 2011. Effects of intraoperative breaks on mental and somatic operator fatigue: a randomized clinical trial. Surg. Endosc. 25 (4), 1245–1250. https://doi.org/10.1007/s00464-010-1350-1.

European Commission, 2015. Commission implementing regulation (EU) 2015/1998 laying down detailed measures for the implementation of the common basic standards on aviation security (Text with EEA relevance). Off. J. Eur. Union L 299/1-142.

Galinsky, T.L., Swanson, N.G., Sauter, S.L., Hurrell, J.J., Schleifer, L.M., 2000. A field study of supplementary rest breaks for data-entry operators. Ergonomics 43 (5), 622–638. https://doi.org/10.1080/001401300184297.

Goh, J., Wiegmann, D.A., Madhavan, P., 2005. Effects of automation failure in a luggage screening task: a comparison between direct and indirect cueing. In: Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society, pp. 492–496 26-30 September, Orlando: FL.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Wiley, New-York, NY.
Gulowsen, J., 1972. A measure of work group autonomy. In: Davis, L.E., Taylor, J.C. (Eds.), Design of Jobs. Penguin, Middlesex, England, pp. 374–390.
Hackman, J.R., Lawler, E.E., 1971. Employee reactions to job characteristics. J. Appl. Psychol. 55 (3), 259–286. https://doi.org/10.1037/h0031152.
Hackman, J.R., Oldham, G.R., 1976. Motivation through the design of work: test of a theory. Organ. Behav. Hum. Perform. 16 (2), 250–279. https://doi.org/10.1016/0030-5073(76)90016-7.
Hallbeck, M.S., Lowndes, B.R., Bingener, J., Abdelrahman, A.M., Yu, D., Bartley, A., Park, A.E., 2017. The impact of intraoperative microbreaks with exercises on surgeons: a multi-center cohort study. Appl. Ergon. 60, 334–341. https://doi.org/10.1016/j.apergo.2016.12.006.
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: Proceeedings of the 39th IEEE Carnahan Conference on Security Technology, pp. 189–192.
Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), Human Mental Workload. North-Holland, Amsterdam, pp. 139–183.
Hättenschwiler, N., Sterchi, Y., Mendes, M., Schwaninger, A., 2018. Automation in airport security X-ray screening of cabin baggage: examining benefits and possible implementations of automated explosives detection. Appl. Ergon. 72, 58–68. https://doi.org/10.1016/j.apergo.2018.05.003.
Helton, W.S., Dember, W.N., Warm, J.S., Matthews, G., 1999. Optimism, pessimism, and false failure feedback: effects on vigilance performance. Curr. Psychol. 18 (4), 311–325. https://doi.org/10.1007/s12144-999-1006-2.
Henning, R.A., Jacques, P., Kissel, G.V., Sullivan, A.B., Alteras-Webb, S.M., 1997. Frequent short rest breaks from computer work: effects on productivity and well-being at two field sites. Ergon 40 (1), 78–91. https://doi.org/10.1080/001401397188396.
Henning, R.A., Kissel, G.V., Maynard, D.C., 1994. Compensatory rest breaks for VDT operators. Int. J. Ind. Ergon. 14 (3), 243–249. https://doi.org/10.1016/0169-8141(94)90100-7.
Henning, R.A., Sauter, S.L., Salvendy, G., Krieg Jr., E.F., 1989. Microbreak length, performance, and stress in a data entry task. Ergon 32 (7), 855–864. https://doi.org/10.1080/00140138908966848.
Inagaki, T., 2003. Adaptive automation: sharing and trading of control. In: Hollnagel, E. (Ed.), Handbook of Cognitive Task Design. Lawrence Erlbaum Publishers, Mahwah, N.J, pp. 221–245.
JASP Team, 2018. JASP (version 0.8.6)[computer software]. https://jasp-stats.org/.
Jian, J.-Y., Bisantz, A.M., Drury, C.G., 2000. Foundations for an empirically determined scale of trust in automated systems. Int. J. Cogn. Ergon. 4 (1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04.
Jones, E.D., 1919. The Administration of Industrial Enterprises. Longman, New-York.
Kaber, D.B., Riley, J.M., 1999. Adaptive automation of a dynamic control task based on secondary task workload measurement. Int. J. Cogn. Ergon. 3 (3), 169–187.
Karasek, R., Theorell, T., 1990. first ed. Healthy Work: Stress Productivity and the Reconstruction of Working Life, vol. XIII. Basic Books, New York, pp. 381 (S. : graph. Darst).
Kirschenbaum, A., 2015. The social foundations of airport security. J. Air Transp. Manag. 48, 34–41. https://doi.org/10.1016/j.jairtraman.2015.06.010.
Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., others, 2007. What's new in Psychtoolbox-3. Perception 36 (14), 1.
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. Ergon 52 (6), 644–656. https://doi.org/10.1080/00140130802526935.
Kopardekar, P., Mital, A., 1994. The effect of different work-rest schedules on fatigue and performance of a simulated directory assistance operator's task. Ergon 37 (10), 1697–1707. https://doi.org/10.1080/00140139408964946.
Kroemer, K.H.E., Grandjean, E., 2000. Fitting the Task to the Human: A Textbook of Occupational Ergonomics, fifth ed. Taylor & Francis, London 416 pp.
Kühnel, J., Zacher, H., Bloom, J. de, Bledow, R., 2016. Take a break!: benefits of sleep and short breaks for daily work engagement. Eur. J. Work Organ. Psychol. 26 (4), 481–491. https://doi.org/10.1080/1359432X.2016.1269750.
Lee, J.D., Moray, N., 1992. Trust, control strategies and allocation of function in human-machine systems. Ergon 35 (10), 1243–1270. https://doi.org/10.1080/00140139208967392.
Lim, J., Kwok, K., 2016. The effects of varying break length on attention and time on task. Hum. Factors 58 (3), 472–481. https://doi.org/10.1177/0018720815617395.
Macmillan, N.A., Creelman, C.D., 2005. Detection Theory: A User's Guide, first ed. Psychology Press, Mahwah, NJ, pp. 492.
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.
McLean, L., Tingley, M., Scott, R.N., Rickards, J., 2001. Computer terminal work and the benefit of microbreaks. Appl. Ergon. 32 (3), 225–237. https://doi.org/10.1016/S0003-6870(00)00071-5.
Mery, D., Mondragon, G., Riffo, V., Zuccar, I., 2013. Detection of regular objects in baggage using multiple X-ray views. Insight 55 (1), 16–20. https://doi.org/10.1784/insi.2012.55.1.16.
Meyer, J., 2001. Effects of warning validity and proximity on responses to warnings. Hum. Factors 43 (4), 563–572. https://doi.org/10.1518/001872001775870395.
Ostarek, M., Joosen, D., Ishag, A., Nijs, M. de, Huettig, F., 2019. Are visual processes causally involved in "perceptual simulation" effects in the sentence-picture verification task? Cognition 182, 84–94. https://doi.org/10.1016/j.cognition.2018.08.017.
Parasuraman, R., Mouloua, M., Molloy, R., Hilburn, B., 1993. Adaptive function allocation reduces performance costs of static automation. In: 7th International Symposium on Aviation Psychology. 7th International Symposium on Aviation Psychology, Columbus. 26-29 April. Ohio State University, Columbus, pp. 178–185.
Parasuraman, R., Sheridan, T.B., Wickens, C.D., 2000. A model for types and levels of human interaction with automation. IEEE Trans. Syst., Man, Cybern. A 30 (3), 286–297. https://doi.org/10.1109/3468.844354.
Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vis. 10 (4), 437–442.
Rice, S., McCarley, J.S., 2011. Effects of response bias and judgment framing on operator use of an automated aid in a target detection task. J. Exp. Psychol. Appl. 17 (4), 320–331. https://doi.org/10.1037/a0024243.
Salthouse, T.A., Somberg, B.L., 1982. Skilled performance: effects of adult age and experience on elementary processes. J. Exp. Psychol. Gen. 111 (2), 176–207. https://doi.org/10.1037/0096-3445.111.2.176.
Sauer, J., Chavaillaz, A., 2018. How operators make use of wide-choice adaptable automation: observations from a series of experimental studies. Theor. Issues Ergon. Sci. 19 (2), 135–155. https://doi.org/10.1080/1463922X.2017.1297866.
Sauer, J., Kao, C.-S., Wastell, D., 2012. A comparison of adaptive and adaptable automation under different levels of environmental stress. Ergon 55 (8), 840–853. https://doi.org/10.1080/00140139.2012.676673.
Scerbo, M.W., 2006. Dynamic function allocation. In: Karwowski, W. (Ed.), International Encyclopedia of Ergonomics and Human Factors, second ed. CRC/Taylor & Francis, Boca Raton, FL, pp. 1080–1082.
Schwaninger, A., Hardmeier, D., Hofer, F., 2005. Aviation security screeners visual abilities & visual knowledge measurement. IEEE Aerosp. Electron. Syst. Mag. 20 (6), 29–35.
Singh, S., Singh, M., 2003. Explosives detection systems (EDS) for aviation security. Signal Process. 83 (1), 31–55. https://doi.org/10.1016/S0165-1684(02)00391-2.
Sluiter, J.K., Frings-Dresen, M.H.W., Meijman, T.F., van der Beek, A.J., 2000. Reactivity and recovery from different types of work measured by catecholamines and cortisol: a systematic literature overview. Occup. Environ. Med. 57 (5), 298–315. https://doi.org/10.1136/oem.57.5.298.
Steinborn, M.B., Huestegge, L., 2016. A walk down the lane gives wings to your brain. Restorative benefits of rest breaks on cognition and self-control. Appl. Cognit. Psychol. 30 (5), 795–805. https://doi.org/10.1002/acp.3255.
Strongman, K.T., Burt, C.D., 2000. Taking breaks from work: an exploratory inquiry. J. Psychol. 134 (3), 229–242. https://doi.org/10.1080/00223980009600864.
Teichner, W.H., 1974. The detection of a simple visual signal as a function of time of watch. Hum. Factors 16 (4), 339–352. https://doi.org/10.1177/001872087401600402.
Tucker, P., 2003. The impact of rest breaks upon accident risk, fatigue and performance: a review. Work. Stress 17 (2), 123–137. https://doi.org/10.1080/0267837031000155949.
Wales, A.W.J., Anderson, C., Jones, K.L., Schwaninger, A., Horne, J.A., 2009. Evaluating the two-component inspection model in a simplified luggage search task. Behav. Res. Methods 41 (3), 937–943. https://doi.org/10.3758/BRM.41.3.937.
Warm, J.S., Parasuraman, R., Matthews, G., 2008. Vigilance requires hard mental work and is stressful. Hum. Factors 50 (3), 433–441. https://doi.org/10.1518/001872008X312152.
Wells, K., Bradley, D.A., 2012. A review of X-ray explosives detection techniques for checked baggage. Appl. Radiat. Isot. 70 (8), 1729–1746. https://doi.org/10.1016/j.apradiso.2012.01.011.