ORIGINAL ARTICLE

Learning game for improving forklift drivers' safety awareness

Esko Lehtonen[1,2] · Pia Perttula[3] · Ida Maasalo[1,4] · Kaisa Reuna[3] · Henriikka Kannisto[3] · Vuokko Puro[3] ·
Maria Hirvonen[3]

Abstract
Forklift drivers in in-house logistics are at a high risk of occupational injuries. This study piloted a video-based learning game for forklift drivers, designed to increase their safety awareness in their highly dynamic operating environment. The game consisted of video clips filmed from the driver's perspective. Each video clip was stopped and masked at a point where a safety-relevant item or hazard appeared. Two locations were marked on the screen, one with the hazard and the other with no hazard. The players had to choose the location with the hazard while the screen was masked and received feedback on their choice. Eighty-one professional forklift drivers played the game at their workplaces during the 3–4 month intervention. In addition, forklift registered hard shocks were recorded and compared to the game performance of 20 participants. Higher accuracy in the game correlated with a lower shock rate, suggesting that real-world forklift driving skills transferred to the game. However, accuracy improvement in the game did not decrease the vehicle registered hard shocks. During the intervention, 40% of the participants reported that they began to pay more attention to the hazards of their work environment. Overall, the results suggest that a gamified approach could be useful for demonstrating safety-relevant features at the workplace and supporting drivers' safety-related self-reflection.

Keywords Digitalization · Safety · Occupational risk · Accident risk · In-house logistics · Situation awareness

1 Introduction

Logistics involve a high risk of occupational injuries. In Finland, material transfers[1] cause one-quarter of all fatal occupational accidents (Perttula and Salminen 2012). Overall, occupational accidents related to material transfers also cause longer term disabilities than other occupational accidents (Perttula et al. 2003). Material transfer-related accidents occur especially in in-house logistics, i.e. in warehouses and terminals (Perttula 2013).

✉ Esko Lehtonen
  esko.lehtonen@vtt.fi

1 Traffic Research Unit, University of Helsinki, Helsinki, Finland

2 VTT Technical Research Centre of Finland, Espoo, Finland

3 Finnish Institute of Occupational Health, Helsinki, Finland

4 Finnish Road Safety Council, Helsinki, Finland

Occupational accidents in in-house logistics have often been studied from the accident prevention perspective. Studies promoting occupational safety in warehouses have highlighted risk factors such as physical strain in both manual and mechanical material transfers (Denis et al. 2006; Lanoie and Trottier 1998), as well as risks caused by the transferred material itself (St-Vincent et al. 2005). The prevention of physical strain has been promoted by ergonomic studies (Roman-Liu 2010; Mack et al. 1995). In addition to the traditional background factors of occupational accidents, some studies of the warehouse sector have also focused on psychological and behavioural factors (de Koster et al. 2011).

Accidents in in-house logistics often involve forklifts (de Koster et al. 2011). Busy traffic flows, different forms of transfers using the same routes and time pressure are often factors in such accidents. To work safely, drivers need to manage their work environment: to be aware of people moving in their work area and to control the movements of their vehicles (Saric et al. 2013). Their confined work environments make this challenging. Visibility is often limited due

---

[1] The process of transferring material from one place to another (manually or mechanically).

to the inadequate lighting and design of the warehouses. In addition, the load is typically located in front of the drivers, further restricting the field of view. Untidiness or inadequate maintenance and markings of forklift routes can also pose a risk. Other risk factors include physical load and psychosocial factors.

Overall, due to limited visibility and a complex, dynamic environment, lack of sufficient situation awareness seems to be a contributing factor to many forklift accidents. To anticipate potential hazards, it is essential to perceive and identify objects and situations that could create a hazard. In terms of Endsley's (1995) classical definition, situation awareness refers to the perception of the relevant elements in the work environment, as well understanding of their meaning and role in the situation, and anticipation of what will happen next. Situation awareness is not only a passive representation of the work environment but also guides drivers' perception, for example, eye movements. In other words, good situation awareness also means effective anticipation of hazards.

Among car drivers, the ability to perceive and anticipate hazardous situations has been identified as one of the few driving skills related to safety (Horswill and McKenna 2004). Novice drivers are involved in crashes much more often than experienced drivers, and they also perform more poorly in tests designed to test hazard perception/anticipation. Studies have been able to link performance in hazard perception/anticipation tests to actual crash involvement (Horswill et al. 2015; Tuske et al. 2018) and occurrences of heavy braking (Hill et al. 2019). Studies of novice and experienced drivers' visual strategies have also suggested that inexperienced drivers' visual search of their environment are more restricted (Crundall and Underwood 1998; Underwood et al. 2002; Lehtonen et al. 2014), as they have less developed schemas or mental models of the situation. In other words, they do not know where they should look. Especially if the hazards are latent (not directly visible), for example, a pedestrian crossing a street obstructed by a parked van (Vlakveld 2014).

Consequently, interactive computer-based programs have been developed to train car drivers' hazard anticipation (McDonald et al. 2015). Many of these programs include some elements of games or utilize game-based learning. Game-based learning has been suggested to activate cognitive processes (Wouters et al. 2013) and provide more motivation for learning than conventional learning methods (Malone and Lepper 1987). Game-based learning can provide immediate feedback to the learner, which is essential for inexperienced drivers in particular, as they are known to overestimate their driving skill (Deery 1999).

1.1 Aims of the current study

In the current study, we implemented a video-based serious game intending to promote forklift drivers' safety awareness, by demonstrating what kinds of hazards may be present in the work environment and directing the drivers to anticipate potential hazards. Hazards were understood to broadly include not only other warehouse users, forklifts and pedestrians, on whom the driver had to keep an eye, but also environmental features such as unsafely placed items.

The game consisted of video clips filmed from the driver's perspective. Each video clip was stopped at a point that presented a potential hazard. The view was masked and two locations were presented to the player, one with the hazard and one with no hazard. The correct answer awarded points and incorrect answers subtracted points. In addition, to support learning, a textual description of the potential hazard was given after answering (Moreno and Mayer 2005). The current game was adapted from a game previously developed to improve situation awareness among cyclists (Lehtonen et al. 2017).

Even though warehouse work involves new technology such as vehicle telemetry, the effect of these new technologies on occupational safety and the fluency of work has not yet been widely studied. In this study, we did not focus on the effects of vehicle telemetry as such, but we used vehicle telemetry to compare performance in the game and in the real world. Vehicle registered hard shocks can be argued to be a proxy for crashes, in a similar fashion as hard braking is used in driving research (Hill et al. 2019).

The first aim was to investigate whether forklift driving performance could be linked to performance in the game, by correlating the shock rate (the number of forklift registered hard shocks per hour) to the accuracy of responses in the game.

The second aim was to investigate learning in the game, by examining the change in the accuracy and response time over time. Learning was also investigated as a function of different hazard types.

The third aim was to investigate whether the learning in the game transferred into the real-world environment, by examining whether the change in accuracy correlated with the change in shock rate.

Playing the game may directly train skills that are relevant for safe forklift driving. In addition to such direct effects, playing the game may also trigger a self-reflection process in the participants (Hatakka et al. 2002). Through self-reflection, the participants start to pay more attention to safety in their workplace, which may eventually lead to behavioural and organizational changes.

Therefore, the fourth aim was to understand the more indirect effects of the game, by assessing drivers' self-reported attitudes toward the game and what benefits they saw in the game.

Finally, we also investigated whether the players' background factors, such as other computer usage or attitude toward new technology, was linked to performance in the game.

2 Methods

2.1 Learning game

The learning game consisted of video clips filmed from the forklift driver's perspective. The player started the video clip and viewed it until a predefined stopping point. At this point, the video was masked with a grey mask, and two locations were presented as circles on the screen. The task of the player was to choose the location in which a potential hazard had been visible on the video before it was masked (for screenshots, see Supplementary Material).

After the player made the choice, the game presented them with feedback, which consisted of points and a textual explanation to provide guidance (Moreno and Mayer 2005). The player scored two points if they answered correctly within 2 s, and one point when they gave a correct answer after 2 s. If the answer was incorrect and given within 2 s, the player received two minus points, and one minus point was given for an incorrect answer given after 2 s. This was to encourage quicker answers based on observing the environment rather than reasoning based on the overall gist of the scene. The scores were displayed alongside the locations, together with textual feedback. The textual feedback always mentioned the potential hazard on the location (e.g. 'forklift', 'pedestrian', 'an obstacle on the driveway'). The total scores were displayed on the screen during the game as well on the opening screen of the game after logging in. The total score per game could not be negative.

2.2 Videos and hazards

Videos were filmed at seven different locations during normal working hours. A GoPro Hero 4 video camera was mounted on a forklift, filming from the drivers' perspective. The exact mounting position depended on the forklift used by the company. The videos were filmed at 1080p (1920 × 1080) with 30 fps in the aspect ratio of 16:9, using 'wide' mode, giving approximately 170° of field of view.

We collected over 100 h of raw video footage. The situations in the video that featured potential hazards were identified by the researchers, who were experts in workplace safety. The five types of hazards were:

1. Forklift: Other moving forklifts visible in the video.
2. Bluespots: *Bluespots* of moving forklifts (a blue spot of light projected onto the floor in the direction in which the forklift is moving). The actual forklift may be obstructed.
3. Environment: Obstacles, unsafely placed objects on shelves, or similar stationary objects.
4. Mirror: Forklifts, pedestrians, or bluespots visible in the mirror.
5. Pedestrian: All pedestrians visible in the video.

For each suitable situation, a video clip was extracted. If the researchers agreed that a clip represented a good example of a potential hazard, it was included in the set of potential clips for the game. All the identifying information, such as people's faces or company logos were blurred.

A time point was selected on the clip when a hazard was clearly visible and when another location had no such hazard. Care was taken to find time points where the empty location could be placed at a comparable height to the target to give no extra clues to the player.

A description of the hazard was written down and used in the game to give feedback. Hazards were also categorized on the basis of the type of target, and how long the target had been visible before the video was stopped and masked (< 1 s, 1–3 s, > 3 s).

Before the actual research, a pilot version of the game was tested by a group of professional forklift drivers and non-drivers. The game interface and the hazard definitions were improved on the basis of their feedback. The final set of clips was also played to a professional forklift driver instructor, and some clips were edited or removed on the basis of their feedback.

2.3 Video sets

The game had nine sets of videos in total. The first set was a tutorial set of four videos, giving examples of different kinds of targets. Three of the sets were tests (test 1, 2, and 3), which were played once at specific times by all the players. The tests were used to track improvement during the intervention. Each test had 20 videos, selected at random but ensuring an equal number of different hazards between the tests (Table 1). The aim was to include at least three videos for each hazard type and hazard visibility combination. For environmental objects, the video footage did not contain hazards that were visible for less than 1 s. Forklift-related hazards were the most common, and therefore, we included more videos with forklifts.

Five sets of videos were for free playing, each consisting of ten videos, and players were free to play them repeatedly. Free playing was possible either between tests 1 and 2 or between tests 2 and 3 (see Sect. 2.6).

Table 1 Total number of different kinds of hazards in the three tests

Hazard type	Hazard visibility before video stopped		
	< 1 s	1–3 s	> 3 s
Forklift	9	9	9
Bluespot	3	3	3
Environment	0	3	3
Mirror	3	3	3
Pedestrian	3	3	3

The hazards were equally distributed into the three tests; thus the number of hazards is always divisible by three

2.4 Forklift telemetry data

Some of the forklifts used by the drivers were connected to a fleet management system (I_Site). The system functions by installing a data collection unit, equipped with a 3G connection, to a forklift, and the data are transferred to an internet portal via a wireless connection. Four target companies committed to releasing the data collected through the fleet management system for use in this research project. The information material was collected using the data collection units installed on 161 forklifts in the companies.

The information obtained from the asset management system was analysed during this research project for data points such as impacts occurring during work shifts, forklift utilization and the times and dates of undesired events. The system logged the information automatically when the forklift was in use. The system required each driver to have a personal code, which they had to use to log in before they could drive the forklift. This personal code was used to compare the forklift driving sessions to the performance in the game. The data from the technology installed on the forklifts were collected for 1.5–2 years in each company. The data collection period started before the game intervention and ended after it.

The vehicle telemetry data were first screened, and the total number of hours driven per driver each month and overall was calculated. Three driver IDs with over 200 h of recorded driving per month were excluded, because it

was likely that these personal forklift IDs had been used by multiple drivers.

The fleet management system had a proprietary algorithm to detect *hard shocks* from the vehicle telemetry data using the accelerator signals. The shock rates were then calculated by dividing the number of shocks by the number of hours driven. Shock rates were calculated for the periods before the game intervention, during, and after, as well as overall, for each driver.

2.5 Participants

Participation in the interventions was voluntary. All the participants gave written informed consent to participate in the study. A total of 108 participants started the game intervention, but only 81 completed it (Table 2).

Age was recorded in only four categories (< 30, 31–40, 41–50, over 50 years) to ensure the anonymity of the participants. All the four age groups were well represented among those who completed all three tests (n = 81): 21% of the participants were less than 30, 35% were between 31 and 40, 23% were between 41 and 50 and 21% were over 50 years old. The forklift truck driving experience of the forklift drivers who began the intervention varied from less than 2 years to over 20 years: 1% of the participants had been driving forklift trucks for less than 2 years, 15% had been driving for 2–5 years, 23% had been driving for 6–10 years and 61% had been driving for 11 years or more. Almost half of the participants (46%) reported that they did not play computer games at all and 22% played computer games weekly. Computer usage varied among the participants: 15% used computers for more than 6 h per day for activities other than games, 60% used them for 1–6 h per day, 15% used them for less than 1 h/day, and 10% never used a computer.

2.6 Procedure

Participation in the intervention was voluntary in each of the co-operating workplaces. Before beginning the learning game, the forklift drivers filled in a preliminary questionnaire which included questions on their previous

Table 2 Number of participants who participated in test and could be used in specific analyses

Description	Number of participants	Analyses performed (section number in parentheses)
Participants who started intervention	108	
Participants who completed all three tests	81	Learning in the game (3.2), self-reported evaluation of game (3.4), analysis of background factors (3.5)
… and had more than 10 h of driving data.	20	Correlation between shock rate and game performance (3.1)
… and had 5 h driving both during the before and after periods.	13	Correlation of change in game performance and change in shock rate (3.3)
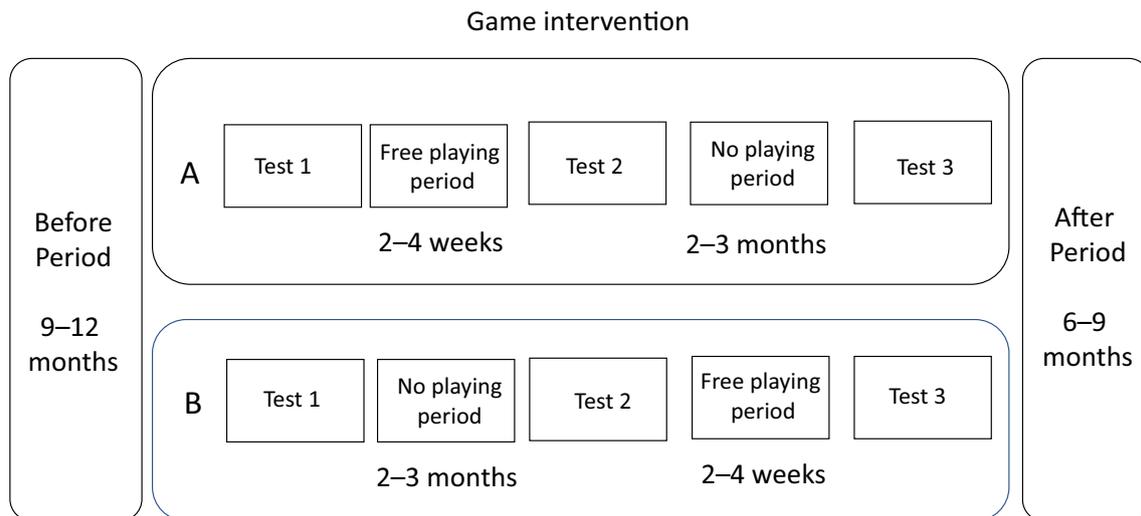
Fig. 1 Schedule of game intervention and before and after periods

experience and activities on computer games. Before the first test, they were given their personal player ID and instructions for the game. Then, they practised with four video clips before beginning Test 1, which they completed independently following the instructions given during the game. Access to the game was granted with a password and individual player ID.

The six workplaces were divided into two categories based on the scheduling of the free playing period within the game intervention (Fig. 1). All the players first played the tutorial set with personal guidance on the game, after which they played test 1. In three workplaces (group A in Fig. 1), test 1 was followed by a free play period, after which test 2 was completed, and test 3 was played by all the players after a period of non-play. In another three workplaces (group B in Fig. 1), the free playing period was between tests 2 and 3. In all six workplaces, the free playing period lasted 2–4 weeks. During this time, the participants were able to play the game on their own. They were sent 2–3 SMS messages to remind them of the ongoing period of free playing. The vehicle telemetry was recorded and stored before, during and after the game intervention. The durations of these before and after periods varied in the workplaces (Fig. 1).

Tests 1, 2 and 3 were performed in attendance of a researcher. The game was not accessible during the no playing periods.

After the learning game intervention finished, the participants completed the final questionnaire which asked them about their experiences of the learning game intervention. The final questionnaire included questions on the physical and psychosocial load of the game, and its easiness and usefulness. The participants also reported whether or not they began paying more attention to their work environment after playing the game.

2.7 Analyses

Performance in the game and the shock rates were compared among the drivers who had completed all three tests, and who also had sufficient available vehicle telemetry data (see Sect. 3.1). The vehicle telemetry from the fleet management system could be linked to 43 of the drivers who completed the intervention. We also excluded driver IDs who had logged less than 10 h of driving a telemetry-equipped forklift. The choice of the cut point was based on the observation that the median shock rate among the 43 drivers was 0.09 hits per hour ($M = 0.20$, $SD = 0.36$, $Min = 0$, $Max = 1.83$). Therefore, within 10 h, it was more probable to get a hit than not. This resulted in data from 20 drivers in which performance in the game could be correlated with the hard shocks (Table 2). These drivers had $M = 340$ h of logged driving ($SD = 478$, $Mdn = 122$, $Min = 11$, $Max = 1710$).

The *overall accuracy in three tests and the overall shock rates were compared* using the data from the aforementioned 20 drivers. We also further analysed the driving data collected only during the game intervention (i.e. excluding the before or after period data) and the accuracy in the third test (14 out of 20), because these shock rate and game performance data were collected closer in time. Before calculating the Pearson correlation, the values were transformed to linearize the relationship. As the accuracy was limited to between 0 and 1 and skewed to the right, it was log10

transformed. Shock rate had a lower bound of zero with a long tail and included zero values. Therefore, we used the square root transformation.

At the end, 13 of the 20 drivers had at least 5 h driving both before and after the game intervention (Table 2). These drivers were used to investigate *whether the change in hard shocks correlated with a change in game performance* (see Sect. 3.3).

We used the data of the 81 drivers who had completed all the 3 tests to investigate *learning in the game* (see Sect. 3.2) (Table 2). We calculated the average accuracies (percentage of correct answers) and average response times of each driver for each test set and over the tests. We used repeated-measures ANOVA with polynomial contrasts to test whether game performance improved from test 1 to test 3. Among those who had played extra clips, we tested whether the free playing led to even greater improvement in test performance by running mixed ANOVAs with test (1, 2, 3) as a within-subject factor and the timing of the free playing (1–2 or 2–3) as a between-subject factor.

In addition, to investigate *the effect of learning as a function of target type and the target visibility* in the game data, we calculated average accuracy (proportion of correct answers) as a function of the video clips presented (including the test and free playing clips). After taking log10 transformation from the number of video clips presented, we fit logistic regressions to the data to investigate the effects of the target type and how long the hazard was visible before the video stopped. The logistic regressions were fitted to the pooled data (ignoring the participant level effects). We used the pooled data, because we had an insufficient number of data points to estimate the influence of the participant. As the data were pooled, analysis of statistical significance in terms of p values was not justified. Instead, we used Approximate Bayesian Computation to fit the models and obtain posterior distribution for the accuracy as a function of the number of video clips played using Python package PyMC3 (version 3.5) (Salvatier et al. 2016). The first logistic regression model included intercepts and slopes for each target type. The second logistics regression model was similar but included intercepts and slopes for each visual preview time.

The same 81 participants were also used when analysing the *self-reported data* (see Sect. 3.4) and *the influence of background factors* (Sect. 3.5).

3 Results

3.1 Performance in game and shock rates

The overall shock rate decreased when the overall accuracy increased, as shown by the negative correlation, $\rho = -0.54$, 95% CI [-0.79, -0.13], $t(18) = -2.7423$, $p = 0.014$)
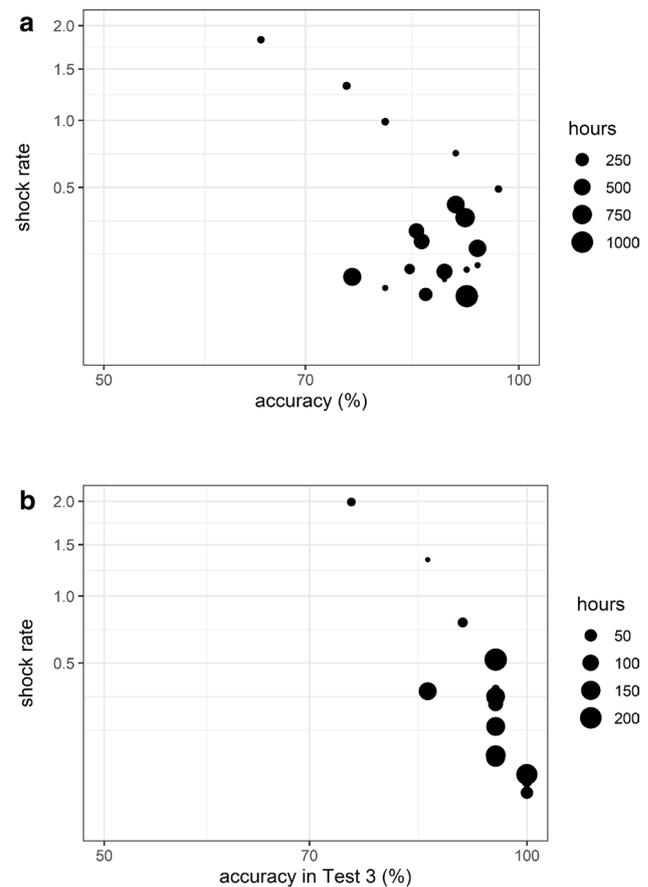
Fig. 2 a Overall accuracy in the game (*x*-axis) and overall shock rate for each player. Total driving hours recorded (during the before and after periods and the game intervention) for driver denoted by size. **b** Accuracy in final test 3 and shock rate during intervention. Total driving hours recorded during game intervention for driver denoted by size. In both figures, the *y*-axis has been square root transformed and the *x*-axis log10 transformed

(Fig. 2a). If we further limit the data to the drivers who drove at least 10 h during the intervention months ($n = 14$), the relationships still holds ($\rho = -0.63$, 95% CI [-0.87, -0.14], $t(12) = -2.783$, $p = 0.017$). It seems that the shock rate during the intervention predicted how high the driver's accuracy would be in the last test, $\rho = -0.89$, 95% CI [-0.96, -0.68], $t(12) = -6.689$, $p < 0.001$) (Fig. 2b).

3.2 Learning in the game

Eighty-one players completed all three tests. Among these, accuracy in the game improved linearly from test 1 to test 3, $F(2, 160) = 67.69$, $p < 0.001$, $\eta_p^2 = 0.46$ (Fig. 3a). In line with accuracy improvement, response times decreased linearly from test 1 to test 3, $F(1.66, 132.80) = 32.87$, $p < 0.001$, $\eta_p^2 = 0.29$, degrees of freedom corrected with Greenhouse–Geisser ($\varepsilon = 0.830$).
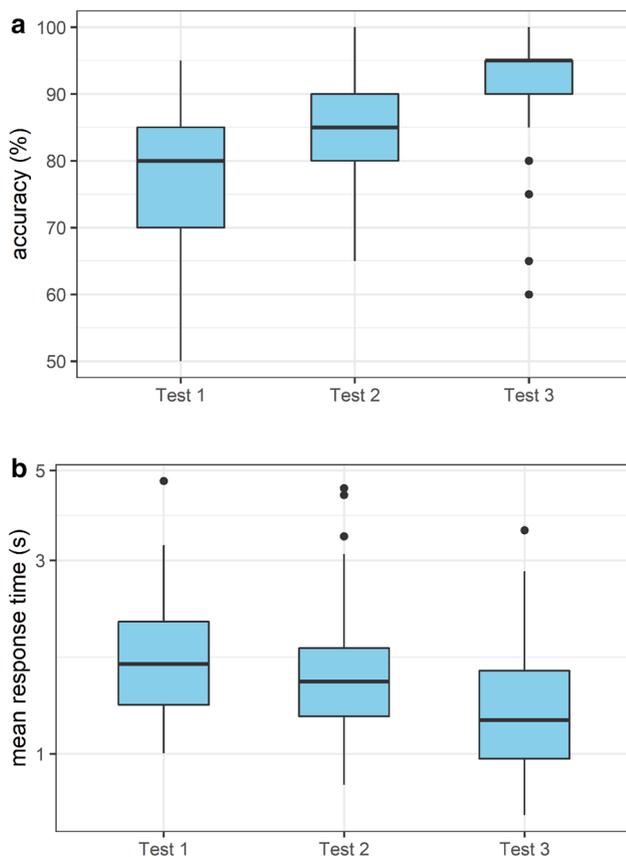
Fig. 3 **a** Players' accuracy in tests. **b** Players' mean response time in tests. *y*-axis is log10 transformed

Thirty-five of the drivers had also played extra clips during their free playing time, between either tests 1 and 2 or tests 2 and 3. The number of extra clips varied from 10 to 160, with $M = 33.95$. Among those who had played extra clips, the timing of the free playing had no significant main effect or interaction ($p > 0.05$).

We also checked whether the drivers who did not perform well in the game were more willing to stop playing and drop out of the intervention. Therefore, we investigated whether the dropouts' accuracy in test 1 was different from that of those who completed all the tests by comparing accuracy among the groups. Simply looking at the means and standard deviation revealed that the groups were not statistically significantly different. The 12 drivers who completed only test 1 had an average accuracy of 75% (SD = 14) in test 1. The 15 drivers who completed both test 1 and test 2 had an average accuracy of 78% (SD = 9) in test 1. The 81 drivers who completed all 3 tests had an average accuracy of 78% (SD = 11) in test 1.

The analysis showed a clear learning effect in the game. In the following, we take a closer look at how the learning occurred as a function of hazard type and visual preview times. Posterior distributions of accuracy were calculated as a function of the clips played.

Pedestrians and bluespots were hardest to detect at the beginning (Fig. 4). However, their accuracy improved during playing. Forklifts were the easiest. How long the hazard was visible for before the video was stopped influenced accuracy. Hazards that were visible for less than 1 s were harder at the beginning than hazards with a longer preview time (Fig. 5).

### 3.3 Correlation of shock rate and accuracy improvement

We also examined whether shock rates decreased during the game intervention, and whether a larger decrease was associated with a larger improvement in accuracy from test 1 to test 3. This analysis was restricted to the subgroup of drivers who had logged at least 5 h of driving before and after the game intervention and who had completed all three tests ($n = 13$). The change in shock rate ($M = 0.01$, SD = 0.30) was not significant, $t(12) = 0.12$, $p = 0.91$, but the change in accuracy was ($M = 13\%$, SD = 9), $t(12) = 5.23$, $p < 0.01$. The change in accuracy from test 1 to test 3, and the change in shock rate from before to after the game intervention were not significantly correlated, $\rho = 0.15$, 95% CI [−0.44, 0.65], $t(11) = 0.49$, $p = 0.63$ (Fig. 6).

### 3.4 Self-reported usefulness and evaluation of the game

The participants who completed all three tests ($n = 81$) were also asked to evaluate the game after the third test. The participants rated the game as non-strenuous physically but felt some time pressure while playing (Table 3). Mostly, the game was evaluated as fun, easy and useful (Table 4). Forty per cent of the participants reported that they began to pay more attention to the risks of their work environment: to things such as other workers, the bluespots of other forklifts and the rear-view mirror.

3.5 Drivers' background and perceptions

We also investigated whether other computer usage and perceptions of new technology were associated with performance in the game or the number of clips played during the game intervention. However, we found no relationships, except for previous experience of computer games. The players who played computer or console games during their free time for more than 6 h/week had shorter response times than those who did not play at all or played only a few times a year (Table 5).
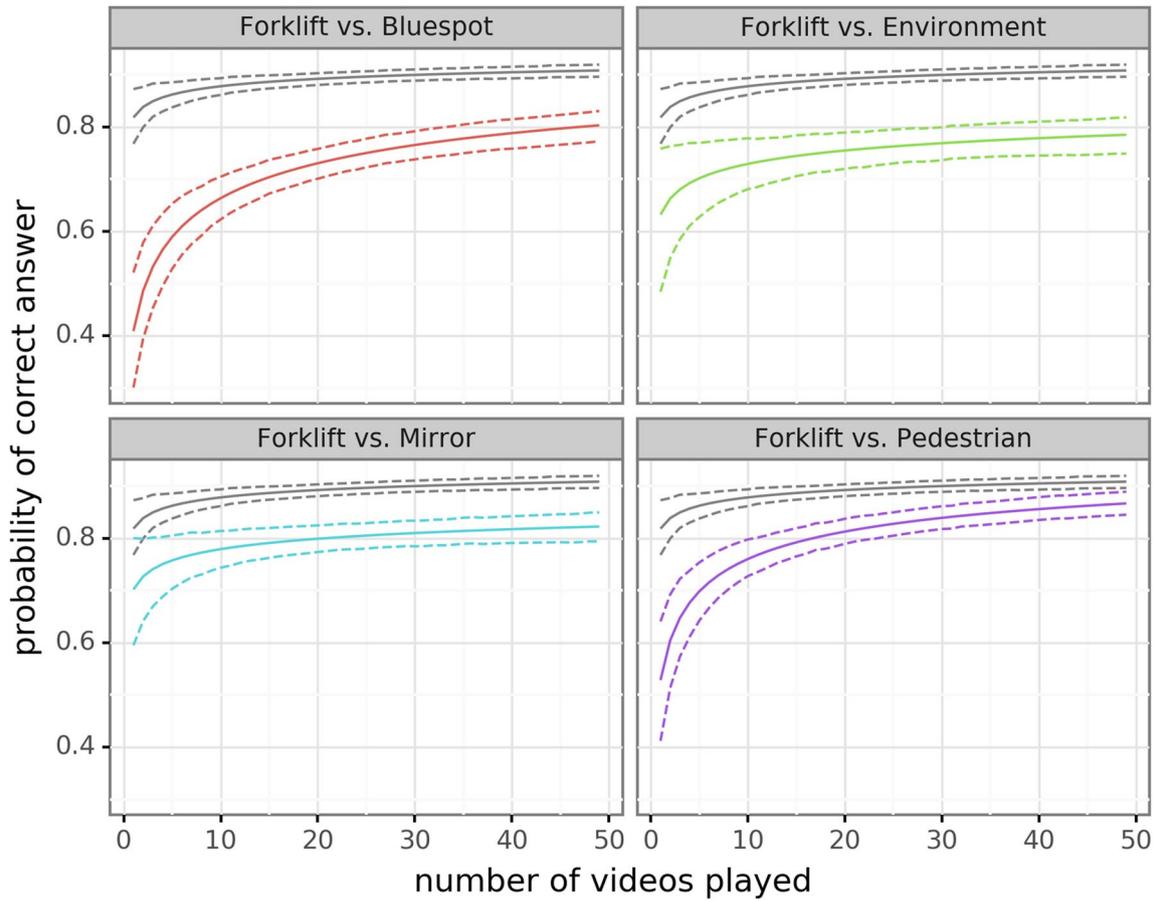
Fig. 4 Learning curves for different target types as a function of videos played. Averages of posterior distribution (solid lines) with 95% highest probability intervals (dashed lines). Values for forklifts are always shown in grey
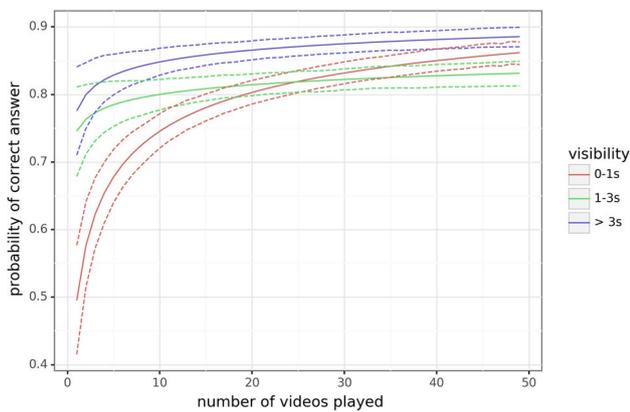
Fig. 5 Learning curves for targets with different visibility times as a function of videos played. Averages of posterior distribution (solid lines) with 95% highest probability intervals (dashed lines)
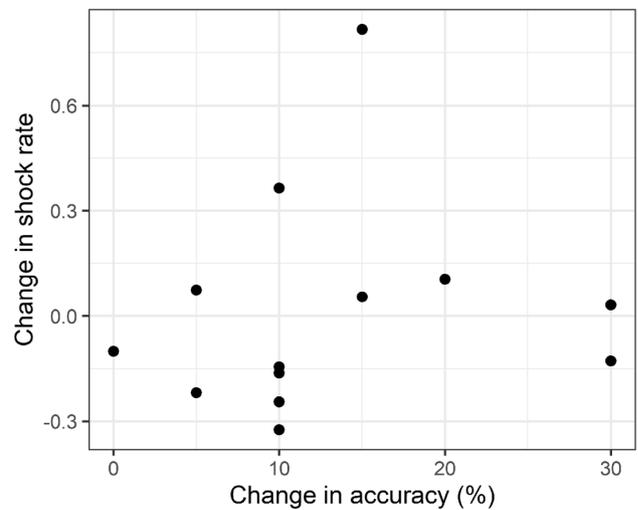
Fig. 6 Change in accuracy in percentage units from first to last test (*x*-axis) and change in shock rate from before phase to after phase (*y*-axis)

Table 3 Participants' self-evaluation of mental workload after test 3 ($n=81$)

Statement	Mean	SD
Estimate the physical strenuousness of playing the game [1 = very low, 10 = very high]	1.93	1.64
Estimate the mental strenuousness of playing the game [1 = very low, 10 = very high]	2.94	2.05
Estimate the feeling of time pressure caused by the game [1 = very low, 10 = very high]	3.86	2.35
How well did you succeed in the game? [1 = perfectly, 10 = totally failed]	4.86	2.51
How hard did you have to strive in order to succeed in the game? [1 = very little, 10 = very much]	4.06	2.33
How insecure, irritated, nervous or distressed did you feel while playing the game? [1 = very little, 10 = very much]	2.93	2.08

Table 4 Participants' opinion of the learning game ($n=81$)

Do the following adjectives describe your opinion of the game	% of the 'yes' answers
Fun	73.6
Easy	64.0
Useful	80.5
Useless	11.8
Boring	17.4
Difficult	17.6

experienced car drivers safer than novices (Deery 1999; Horswill and McKenna 2004).

Drivers' performance in the game improved during the game intervention, measured in terms of accuracy and response times. This suggests that the players learnt something relevant to game performance while playing. However, improvement in the game was not associated with a decrease in vehicle telemetry registered hard shocks, and thus it is not possible to conclude that the game improved drivers' safety. Based on the aforementioned results, we can interpret that performance in the game relies at least partly on the players' situation awareness of the driving

Table 5 Drivers' computer-game playing activity and their accuracy and mean response times ($n=81$)

Do you currently play computer games or console games?	N	Mean accuracy	Accuracy, 95% CI	Average response time (s)	Response time, 95% CI
Not at all/almost never	37	74%	[70, 79]	1.78	[1.56, 1.99]
A few times a year	18	80%	[71, 89]	1.53	[1.35, 1.71]
1–3 h a month	8	79%	[69, 90]	1.46	[1.12, 1.80]
1–3 h a week	6	73%	[50, 95]	1.55	[1.01, 2.08]
4–6 h a week	7	87%	[76, 98]	1.34	[0.93, 1.75]
More than 6 h a week	5	81%	[78, 84]	1.14	[1.02, 1.27]

4 Discussion

We developed a learning game for forklift drivers to improve their safety awareness at the workplace, by demonstrating what kind of hazards can be present in the work environment and guiding them to anticipate potential hazards. The forklift drivers' performance in the game was compared to vehicle telemetry registered shocks, which were used as a surrogate measure of safety.

Drivers who were accurate in the game from the start also had a lower shock rate. This finding is consistent with studies conducted among car drivers. Better hazard anticipation skills have been linked to fewer hard brakings (Hill et al. 2019) as well as fewer crashes (Tuske et al. 2018). Hazard anticipation is a form of situation awareness, and it is one of the most important skills which makes, for example,

situation presented in the video, as the drivers were able to transfer their real-world hazard anticipation skills to the game, and because the drivers who had less shocks scored higher in the game. On the other hand, the skills used in playing the game may not directly transfer to the real-world environment.

There are at least two possible reasons for the lack of transfer from game performance to hard shocks. First, the skills acquired in the game may be too specific to be useful. Second, we only had a very limited amount of driving data available, especially for evaluating the change before and after the game intervention. We also found that the active computer game players had shorter response times in the game than the non-players, which indicates that the test result may reflect prior computer use rather

than the skills relevant for the forklift driving itself. This is important to remember when interpreting any results of a computerized test.

The participants' self-reported evaluation of the game suggests that the game may have had an indirect positive impact. The players evaluated the game mostly positively and assessed that it could be especially useful for novice forklift drivers. Furthermore, previous analysis of the initial questionnaire at the beginning of the intervention suggested that the forklift drivers generally saw the new technical solutions as positive (Perttula et al. 2017).

Forty percent of the participants reported that they began to pay more attention to the risks of their work environment. They specifically mentioned paying more attention to other workers, the bluespots of other forklifts, and their rear-view mirrors. The same hazard types were also hardest for the players at the beginning of the game. Hazards that were visible for less than 1 s before the question in the game were also very hard for the players at the beginning of the game, but their learning effect was also strongest. This suggests that drivers may have learnt to use some indirect cues present in the video to anticipate the appearance of these hazards.

We can interpret the drivers' self-reports as indicating that playing the game engaged them in self-reflection on the possible hazards in their workplace. This self-reflection process can lead to real-world behavioural changes and improve safety in the long term (Hatakka et al. 2002). The benefit of this gamified approach may be that it can trigger such self-reflection processes in a way that the forklift drivers find fun and useful.

However, it is important to remember that the current driver-focused, game-based approach is only one way to improve workplace safety. For example, workplaces can be designed to minimize interactions between forklifts and pedestrians (Horberry et al. 2004). Drivers' situation awareness in particular could be supported by improving the visibility, lighting and routing of work environments (e.g. Saric et al. 2013).

5 Limitations and future studies

The main limitations of the current study are that only a limited amount of driving data was available. This made it more difficult to identify transfer from the game, and did not enable us to control for other possibly influencing factors, such as the drivers' backgrounds and workplaces. As a surrogate measure for safety, registered hard shocks can be argued to have good face validity: if there were no hard shocks, driving would be smooth and no crashes would occur. However, it is important to note that shock rates measured from the vehicle telemetry do not only measure drivers' hazard anticipation skills; they are also strongly influenced by the

work environment, tasks and vehicles used. In the current data, the shock rate levels varied between the workplaces, which may be explained by the differences in the vehicles used, available space, lighting, tidiness, in-house traffic management and the conditions of routes. It is also possible to interpret shock rate as an indicator of an individual driving style. From this perspective, shock rate can be influenced by the driver's cognitive and personality characteristics, in addition to technological and socio-cultural factors and skills for acquiring situation awareness (Sagberg et al. 2015). Further studies should investigate hard shocks in more detail using a larger dataset. The link between hard shocks and accidents should be established on the basis of data, and the same applies to game performance.

Further studies could examine game-based tools as pedagogical learning support and how they function with different pedagogical methods (e.g. flipped-learning, problem-based learning) (cf. Vanderhaegen 2019). The current game was based on videos, which limits the number of situations that can be presented. Immersive virtual-reality technology could also provide richer experience (cf. Yuen et al. 2010).

6 Conclusion

Better performance in the game was linked to fewer vehicle registered hard shocks, but improved performance in the game could not be linked to a decrease in hard shocks. The participants evaluated the game positively and reported that they had started to pay more attention to potential hazards in their workplace. This suggests that the game has the potential to improve safety in in-house logistics and could be used as learning material in forklift driving education as a way to activate learners and engage them in self-reflection. However, the current results did not yet provide conclusive evidence of the benefits of the game.

References

Crundall DE, Underwood G (1998) Effects of experience and processing demands on visual information acquisition in drivers. Ergonomics 41(4):448–458. https://doi.org/10.1080/001401398186937

de Koster R, Stam D, Balk B (2011) Accidents happen: the influence of safety-specific transformational leadership, safety consciousness, and hazard reducing systems on warehouse accidents. J Oper Manag 29:753–765

Deery H (1999) Hazard and risk perception among young novice drivers. J Saf Res 30(4):225–236. https://doi.org/10.1016/S0022-4375(99)00018-3

Denis D, St-Vincent M, Imbeau D, Trudeau R (2006) Stock management influence on manual materials handling in two warehouse superstores. Int J Ind Ergon 36:191–201

Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Hum Factors 37(1):32–64. https://doi.org/10.1518/001872095779049543

Hatakka M, Keskinen E, Gregersen NP, Glad A, Hernetkoski K (2002) From control of the vehicle to personal self-control; broadening the perspectives to driver education. Transp Res Part F Traffic Psychol Behav 5(3):201–215

Hill A, Horswill MS, Whiting J, Watson MO (2019) Computer-based hazard perception test scores are associated with the frequency of heavy braking in everyday driving. Accid Anal Prev 122(March 2018):207–214. https://doi.org/10.1016/j.aap.2018.08.030

Horberry T, Larsson TJ, Johnston I, Lambert J (2004) Forklift safety, traffic engineering and intelligent transport systems: a case study. Appl Ergon 35(6):575–581. https://doi.org/10.1016/j.apergo.2004.05.004

Horswill MS, McKenna FP (2004) Drivers' hazard perception ability: Situation awareness on the road. In: Banbury S, Tremblay S (eds) A cognitive approach to situation awareness. Ashgate, Aldershot, pp 155–175

Horswill MS, Hill A, Wetton M (2015) Can a video-based hazard perception test used for driver licensing predict crash involvement? Accid Anal Prev 82:213–219. https://doi.org/10.1016/j.aap.2015.05.019

Lanoie P, Trottier L (1998) Costs and benefits of preventing workplace accidents: going from a mechanical to a manual handling system. J Saf Res 29(2):65–75

Lehtonen E, Lappi O, Koirikivi I, Summala H (2014) Effect of driving experience on anticipatory look-ahead fixations in real curve driving. Accid Anal Prev 70:195–208. https://doi.org/10.1016/j.aap.2014.04.002

Lehtonen E, Airaksinen J, Kanerva K, Rissanen A, Ränninranta R, Åberg V (2017) Game-based situation awareness training for child and adult cyclists. R Soc Open Sci 4(3):160823. https://doi.org/10.1098/rsos.160823

Mack K, Haslegrave C, Gray M (1995) Usability of manual handling aids for transporting materials. Appl Ergon 26(5):353–364

Malone TW, Lepper MR (1987) Making learning fun: a taxonomy of intrinsic motivations for learning. In: Snow RE, Farr MJ (eds) Aptitude, learning and instruction, vol 3. Cognitive and affective process analyses. Lawrence Erlbaum Associates, London, pp 223–253

McDonald CC, Goodwin AH, Pradhan AK, Romoser MRE, Williams AF (2015) A review of hazard anticipation training programs for young drivers. J Adolesc Health 57(1):S15–S23. https://doi.org/10.1016/j.jadohealth.2015.02.013

Moreno R, Mayer RE (2005) Role of guidance, reflection, and interactivity in an agent-based multimedia game. J Educ Psychol 97(1):117–128. https://doi.org/10.1037/0022-0663.97.1.117

Perttula P (2013) In-house transport and handling. OSH Wiki. http://oshwiki.eu/wiki/In-house_transport_and_handling. Accessed 2 Dec 2019

Perttula P, Salminen S (2012) Workplace accidents in materials transfer in Finland. J Occup Saf Ergon 18(4):541–548

Perttula P, Kiurula M, Merjama J, Laitinen H (2003) Accidents in materials handling at construction sites. Constr Manag Econ 21(7):729–736

Perttula P, Ratilainen H, Puro V (2017) Improving occupational safety in in-house logistics with the aid of digital measures. In: Bernatik A, Kocurkova L, Jørgensen K (eds) Proceedings of the 9th international conference on the prevention of accidents at work (WOS 2017). Taylor & Francis Group, London, pp 141–145

Roman-Liu D (2010) Work-related activities: rules and methods for assessment. In: Koradecka D (ed) Handbook of occupational safety and health. CRC Press, Boca Raton, pp 483–496

Sagberg F, Selpi, Piccinini GFB, Engström J (2015) A review of research on driving styles and road safety. Hum Factors J Hum Factors Ergon Soc 57(7):1248–1275. https://doi.org/10.1177/0018720815591313

Salvatier J, Wiecki TV, Fonnesbeck C (2016) Probabilistic programming in Python using PyMC3. PeerJ Comput Sci 2:e55. https://doi.org/10.7717/peerj-cs.55

Saric S, Bab-Hadiashar A, Hoseinnezhad R, Hocking I (2013) Analysis of forklift accident trends within Victorian industry (Australia). Saf Sci 60:176–184

St-Vincent M, Denis D, Imbeau D, Laberge M (2005) Work factors affecting manual materials handling in a warehouse superstore. Int J Ind Ergon 35:33–46

Tūskė V, Šeibokaitė L, Endriulaitienė A, Lehtonen E (2018) Hazard perception test development for Lithuanian drivers. IATSS Res. https://doi.org/10.1016/j.iatssr.2018.10.001

Underwood G, Chapman P, Bowden K, Crundall D (2002) Visual search while driving: skill and awareness during inspection of the scene. Transp Res Part F Traffic Psychol Behav 5(2):87–97. https://doi.org/10.1016/S1369-8478(02)00008-6

Vanderhaegen F (2019) Pedagogical learning supports based on human–systems inclusion applied to rail flow control. Cogn Technol Work. https://doi.org/10.1007/s10111-019-00602-2 (0123456789)

Vlakveld WP (2014) A comparative study of two desktop hazard perception tasks suitable for mass testing in which scores are not based on response latencies. Transp Res Part F Traffic Psychol Behav 22:218–231. https://doi.org/10.1016/j.trf.2013.12.013

Wouters P, Van Nimwegen C, Van Oostendorp H, Van Der Spek ED (2013) A meta-analysis of the cognitive and motivational effects of serious games. J Educ Psychol 105(2):249–265. https://doi.org/10.1037/a0031311

Yuen KK, Choi SH, Yang XB (2010) A full-immersive CAVE-based VR simulation system of forklift truck operations for safety training. Comput Aided Des Appl 7(2):235–245. https://doi.org/10.3722/cadaps.2010.235-245