CrossMark

Exploring Determinants of Pre-movement Delays in a Virtual Crowd Evacuation Experiment

Nikolai W. F. Bode* iD *, Department of Engineering Mathematics, University of Bristol, Bristol BS8 1UB, UK*

Edward A. Codling iD *, Department of Mathematical Sciences, University of Essex, Colchester CO4 3SQ, UK*

Abstract. Understanding evacuations of high-occupancy buildings presents a major challenge in fire safety science. The total time individuals require to exit a building includes the time it takes them to respond to an alarm and decide to evacuate (pre-movement) and the time it takes them to walk along their chosen exit route (movement). Previous work has shown that variation in pre-movement times is responsible for substantial evacuation delays, but few controlled experiments on this have been conducted. Here, we present a virtual experiment that investigates the level of risk individuals take by collecting virtual objects before evacuating. We determine how over 1200 participants, who were recruited from visitors to the London Science Museum, respond to three factors: a reduction in their knowledge of a building, a change in the behaviour of other simulated evacuees and a change in how they are attached to the objects they can collect (potential gain versus loss). We confirm that collecting more objects is risky, as it affects evacuation success. In our experiment, 44.6% of participants choose extreme strategies by collecting either all or none of the available objects before evacuating. While the adoption of extreme strategies is affected by all three factors we investigate, the only factor that significantly increases the average number of objects participants collect, regardless of extreme strategies, is loss aversion. Our work shows the potential of virtual experiments to safely, quickly and cheaply scope processes causing pre-movement time delays in crowd evacuations. This provides a starting point for further research.

1. Introduction

An important aspect of fire safety practice is to facilitate the fast and safe evacuation of high-occupancy buildings. This requires infrastructure solutions, such as the provision of exit routes. Crucially, it also relies on understanding human behaviour in fire emergency evacuations. For example, the design of fire alarms or

---

* Correspondence should be addressed to: Nikolai W. F. Bode, E-mail: nikolai.bode@bristol.ac.uk

exit signs should be tested for their efficiency in alerting or guiding evacuees (e.g. [1, 2]) and the safe passage of pedestrian crowds through evacuation routes should be guaranteed (e.g. [3, 4]). Fire evacuations are typically studied in terms of two main components [5]: the time delay before movement towards an exit is started (pre-movement time) and the time taken to travel to safety (movement time). To better understand the second component, the motion of pedestrian crowds and the movement decision of individuals within them have been studied extensively (see [4] for a review). However, observations in real emergencies and fire drills have shown that individual behaviour in the pre-movement phase is responsible for substantial delays in evacuation times that can exceed the movement time [5–11]. For example, in unannounced fire drills from large retail stores, the pre-movement time contributed 30-50% to the total evacuation time [9] and a survey of evacuees after a fire in a multi-storey office building indicated a broad range of pre-movement times, reaching values of over 10 min [6].

A variety of causes for pre-movement delays in evacuations have been identified (see e.g. [12]). They include difficulties in hearing fire alarms, attempts to obtain more information and other deliberate preparatory actions, such as collecting personal belongings, putting on warm clothes or waiting for others [5–11]. How and when these different aspects affect evacuation times depends on a variety of factors including warning systems, occupant characteristics, building use, training of any staff involved, time of day and type of building [5–11]. Despite this complexity, it has been suggested that pre-movement times are predictable, based on observational data showing that their mean and variation can be estimated for given contexts [13]. To make predictions, researchers have developed theoretical models based on decision-theory or have used empirically determined distributions to compute pre-movement times [2, 14–17]. In this paper, we focus on individual decision-making and how pre-movement delays may be risked if individuals choose to collect objects before they attempt to move to safety. Collecting objects prior to leaving is not a necessary step in evacuations and such behaviour should thus ideally be prevented with the use of adequate training or other interventions.

Many authors report that individuals stop to collect objects, such as keys, wallets, electronic items or clothes before evacuating in fire drills and even in real incidents [5, 6, 10, 18, 19]. This is an acknowledged problem, as indicated by the recommendation for airlines to include into passenger safety briefings the instruction to leave all carry-on baggage behind during an evacuation [20]. However, few authors have investigated what causes individuals to potentially take this risk. One exception is an evacuation experiment with university students that revealed a positive correlation between the degree to which individuals are attached to their belongings and their pre-movement times in fire drills [18]. In this contribution, we use a virtual experiment to undertake a systematic investigation of how the level of risk individuals take by collecting objects before evacuating is influenced by three factors: a reduction in evacuees' knowledge of a building, a change in the behaviour of other evacuees and a change in how evacuees are attached to the objects they can collect (potential gain versus loss).

Different techniques have been used to study pre-movement delays in fire evacuations. Early work focusses on questionnaires and observations after real incidents

or fire drills [5, 9], but increasingly researchers turn to controlled experiments [21] and computer simulations [15]. An additional approach that has been growing in popularity, and that we employ here, is to conduct fully controlled experiments with human participants using a virtual experiment (see [22] for a review). This experimental paradigm has been used widely in evacuation research to study the efficacy of exit signage [1], the occurrence of helping behaviour [23], social influences on movement [24], or exit route choice [12, 25], to name a few examples. Virtual experiments allow unparalleled control over the experimental setting, they are often time and cost effective and they make it possible to study stressful scenarios that cannot be simulated in other experiments for ethical reasons [22, 26]. However, it has rightly been recognised that great care needs to be taken when interpreting findings, as behaviour in virtual environments may not extend directly to real-world contexts [22, 26]. In other words, the ecological validity of experiments in virtual environments is typically untested and any interpretation or use of findings needs to consider this. We cover this issue in more detail in the Discussion.

This paper addresses two main objectives. First, our experiment provides insights that may potentially lead to an increased understanding of risk-taking related to collecting objects in the pre-movement stage of evacuations. Second, our contribution showcases the potential of simple virtual experiments and we make suggestions for how this technology and findings from it could be used to explore other evacuation and fire safety scenarios.

2. Methods

2.1. Experiment

Our experiment uses an established virtual experiment framework to test decision-making and behavioural responses of individual human participants in simulated crowd evacuations [23, 26, 27]. Participants are presented with a top-down view of a building floor plan that contains one simulated pedestrian that they can control via mouse clicks, as well as seventy-four computer-controlled pedestrians that are evenly distributed across the entire environment (Fig. 1). The virtual environment comprises a central room with two doors that are connected via corridors to a single exit from the building (Fig. 1). In addition, the central room contains ten objects participants can collect (yellow disks or 'coins' in Fig. 1). Participants can see the layout of the virtual environment, but they can only see the content of the corridor or room in which they are located. For example, in Fig. 1 the content of the central room is visible, but the content of the corridors is hidden and coloured in grey. In our experiment, we assess the level of risk participants take by recording how many coins they collect prior to leaving the virtual environment after an evacuation is initiated. Our experiment consists of a training phase and a test phase that we describe below. We investigate how three factors affect the level of risk participants take by using different controlled experimental conditions (also explained below). Full details on the underlying pedestrian movement simulation are presented in previous work [27]. Briefly, the movement of pedestrians is mod-
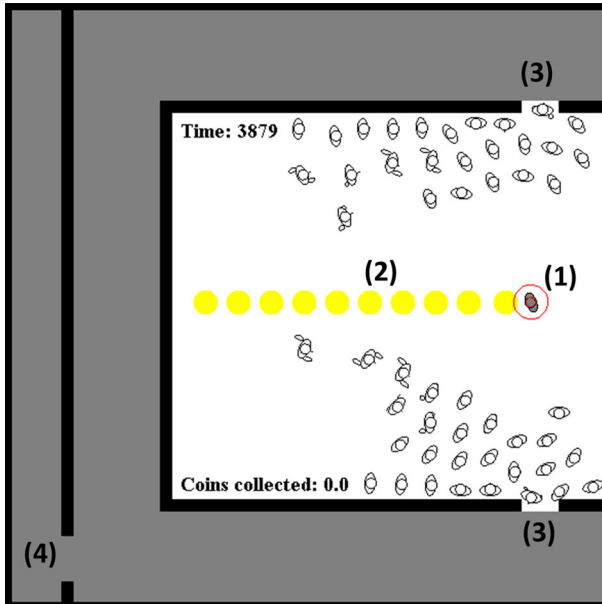
Figure 1. Snapshot of simulated virtual environment at the start of the evacuation task. Participants steer the pedestrian marked in grey, labelled (1). They can collect a maximum of ten coins (2) before exiting the central room through either of the two exits (3) and moving to the final target exit (4). Simulated pedestrians shown in white exit the central room at the start of the evacuation task. The remaining time units and the number of coins collected are indicated in the top-left and bottom-left corners of the central room, respectively. Supplementary video 1 shows a video of the virtual environment.

elled in continuous space. Interactions between pedestrians and between pedestrians and walls are implemented via force vectors acting on simulated pedestrians. The way individuals move towards targets and respond to the built environment (e.g. walls) is encoded in a discrete floor field.

The training phase of the experiment is designed to familiarise participants with the virtual environment and controls, to train them in steering the pedestrian they control and to introduce the coins to them. It is composed of three separate training rounds that are followed by the test phase (Fig. 2). We use popup windows to provide contextual information and instructions as participants advance through the experiment. The only verbal instructions provided in addition to these written messages relate to moving in the virtual environment (e.g. "This is you and if you click somewhere, you will move there in a straight line".). Supplementary video S1 shows a complete run through the experiment including all instructions provided and the length of time they are shown for. In the first training round, participants practice steering the pedestrian they control by following the direction of arrows
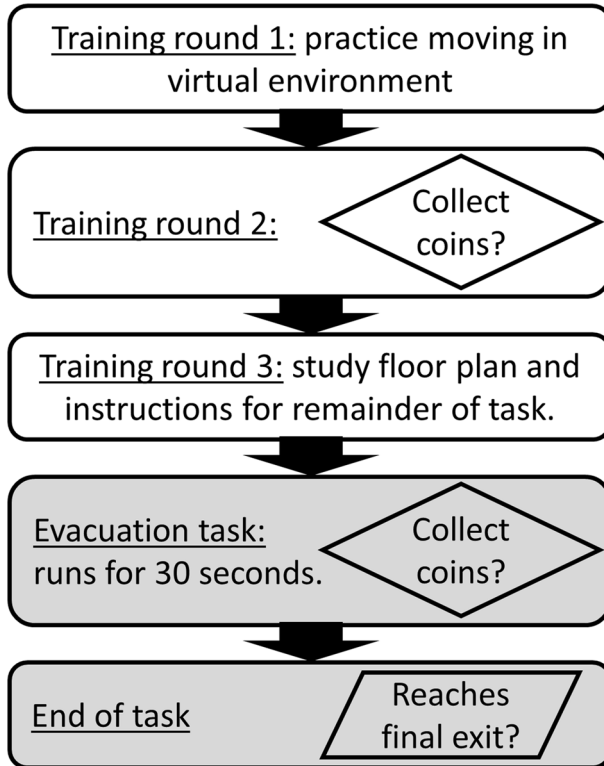
Figure 2. Schematic flow chart of the training phase (white boxes) and test phase (grey boxes) of the virtual experiment.

from a starting location to a target location. In the second training round, participants are introduced to the concept of collecting coins and are alerted to a timer that counts down the time available for the task. They receive the following instructions: "You can now collect money by walking onto the yellow disks. But the available time counts down (top left corner)". Two coins are displayed in this training round. Whenever participants collect a coin, they cannot perform any other actions for 2 s while the timer continues. In this way, we ensure that there is a time cost to collect coins. In the third training round, participants are instructed to look for exits on the floor plan. The location of the final exit is highlighted with a green exit sign. At the end of this round, participants are informed that this is the end of training. They are reminded of the coins and the limited time available: "This was the last training round. Now you have to exit the building. You can still collect coins. The more you collect, the better. BUT, you only have a limited time to evacuate and will lose all coins if you don't evacuate".

The test phase of the experiment is initiated immediately after the end of the training phase with a message that shows a standard green emergency exit sign and reads: "There has been an accident. Leave the building!" The popup window

is displayed for 5 s; immediately after the window closes the ten available coins are then displayed and the computer-controlled pedestrians start to leave the virtual environment, as shown in Fig. 1. Participants must decide how many, if any, coins to collect before leaving the virtual environment through the final exit. The time available to evacuate is 4000 simulated time units which equates to 45 s. This time limit is set to make it difficult, but possible, to collect all coins and evacuate in time. Figure 1 shows the final exit in the bottom-left corner of the virtual environment. For each participant, we randomly place this exit in the top or bottom left corner and we also randomly alternate the direction of the arrows in the first training round to avoid introducing a directional bias in our experiment.

From our virtual experiment, we collected data on the age and gender of participants, as well as the on-screen locations of all simulated pedestrians over time, the mouse click locations of participants, the number of coins collected by participants and whether participants evacuate in time.

2.2. Experimental Treatments

Each participant completed our experiment once. To test what might affect the level of risk-taking by participants (in terms of the number of coins they attempt to collect), we use several different experimental conditions. The scenario described in Sect. 2.1 and shown in Fig. 1 is our baseline or control condition. We implement three further experimental treatments that present participants with scenarios that deviate from the control condition. Each experimental treatment is inspired by realistic factors that could influence the behaviour of evacuees.

The first treatment considers the scenario when people are not familiar with the building layout (labelled treatment 'V'). This situation has previously been identified to cause delays in evacuations [5]. We implement this treatment by not revealing the full floor plan and location of the final exit from the virtual environment in the third training round (see Fig. 3a and compare to Fig. 1).

With the second treatment, we investigate how participants respond to changes in the movement of the simulated crowd, inspired by the established notion of evacuees responding to the behaviour of others (e.g. social influence [2, 10]). In this treatment, labelled 'C', the evacuation of the simulated crowd is delayed and only occurs half-way through the available time (Fig. 3b).

Previous work has suggested that the level of attachment to objects is important in evacuations [18]. To extend this concept, we investigate the widely-studied phenomenon of loss-aversion in human decision-making in our final treatment (labelled 'L'). Prospect theory suggests that people are less sensitive to the possibility of gaining objects or money than to the possibility of losing the same objects [28]. We study the effect of this loss aversion on the level of risk participants are willing to take by changing the instructions relating to the coins that participants can collect. In the second training round, the instructions given in treatment L are: "You have accidentally dropped two coins. You can recover them by walking onto the yellow disks". After the last training round, the instructions are changed to: "This was the last training round. Now you have to exit the building. You have accidentally dropped 10 coins. The more you can recover, the better! BUT,
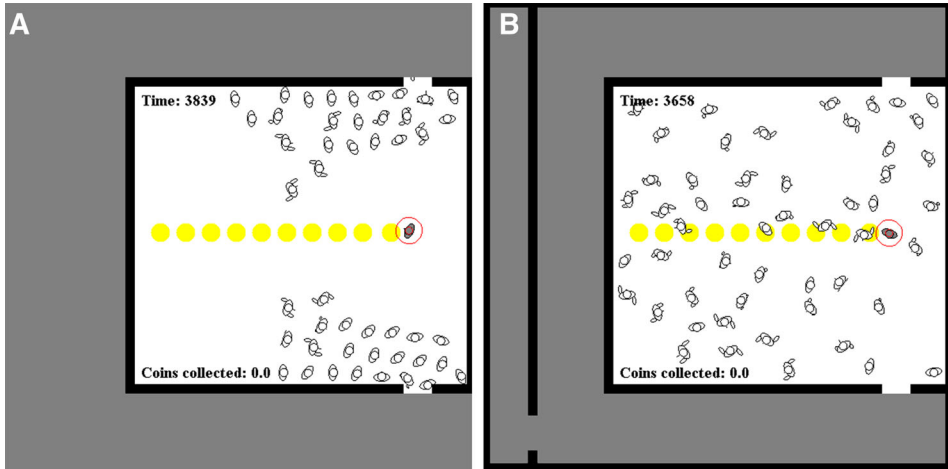
Figure 3.   Simulated virtual environment close to the start of the evacuation for different experimental treatments (the time counter starts at 4000 time units). (a) Treatment V in which the final target exit is not visible and (b) treatment C in which the simulated crowd does not exit straight away, but only half-way through the evacuation task (2000 time units after the evacuation is initiated).

you only have a limited time to evacuate and will lose all coins if you don't evacuate''. Note that all coin rewards are entirely virtual; participants receive neither performance-related nor general monetary or other rewards for taking part in our experiment.

By considering the control condition and all possible combinations of treatments, we obtain a total of eight possible experimental conditions (i.e. control, V, C, L, V + C, V + L, C + L and V + C + L, where 'V + C' implies that both treatments V and C are applied). To ensure a balanced distribution of participants across each of the experimental conditions, we select the condition for each participant in the following way. We assign a unique integer number to each participant and increment this number by 1 between consecutive participants (e.g. the first participant has number 1, the second has number 2, etc.). We label the eight experimental conditions 0 to 7 and assign participants to the condition with the label equal to modulo 8 of their unique number (e.g. participant number 8 implies condition 0). Participants are not allowed to watch others before taking part in the experiment or to talk to others who have already completed the experiment and are therefore not aware of this procedure.

2.3. *Data Collection*

During a public engagement project at the Science Museum in London that ran between the 2nd and 11th of September 2014, we recruited a total of 1229 museum visitors as participants for our experiment. We removed the data of participants that were too young to complete the experiment without the help of their

parents (9 participants younger than 7 years—age was recorded anonymously via the computer program and only accessed during analysis). A total of 139 participants did not report either their gender or their age or both. The data from these participants is not included in statistical analysis that requires this information, but we do include it in all other analysis. Reported genders are split evenly between 597 men (50.8%) and 578 women (49.2%). The first quartile, median and third quartile of reported ages were 20, 23 and 29 years, respectively (mean 25.95 years). Ages ranged from 7 years old to 77. We report the distribution of participants over experimental conditions below.

2.4. Statistical Analysis

We perform three separate statistical analyses on our data in the R programming environment, version 3.2.2 [29].

First, we seek to confirm that the number of collected coins is a determinant for risk-taking in our experiment. Rather than investigating evacuation times that are strongly affected by the virtual experiment design (limited time available), we test the relationship between the binary measure of whether participants evacuated within the time limit (response variable) and the number of collected coins (explanatory variable) using logistic regression. Details for this analysis are included in Appendix A. We use a Likelihood-ratio test to assess if the number of coins participants collect has a non-zero effect on their evacuation success. The Likelihood-ratio test is an appropriate hypothesis test to assess if the fit of a statistical model, measured via the likelihood, improves significantly when an additional parameter is included in the model. After this initial test, we investigate how participant risk-taking behaviour varies across experimental treatments, as follows.

The second analysis of our data is motivated by the observation that many participants adopted extreme strategies by collecting all or no coins. We use logistic regression to investigate how the occurrence of extreme coin collection strategies depends on our experimental treatments, as well as age and gender. Details for this analysis can be found in Appendix B. To avoid having to fit multiple statistical models, as required for the Likelihood-ratio test, we use Wald tests to assess if parameters associated with explanatory variables have a non-zero effect. The Wald test assesses the same null hypothesis as the Likelihood-ratio test and for infinite sample sizes the two tests are asymptotically equivalent. Considering the substantial number of data points in our analysis, the Wald test will therefore provide negligibly different results to the Likelihood-ratio test.

Finally, we investigate how the total number of coins collected by participants, the main numerical measure for the level of risk participants take, depends on the experimental treatments, as well as gender and age. This analysis establishes how participants' risk-taking behaviour depends on the three factors included in our experiment. We use negative binomial regression for this analysis and include additional details in Appendix C. We initially perform this analysis on all data, including participants who adopt extreme strategies (number of coins collected between 0 and 10). Subsequently, to establish the influence of extreme strategy adoption on our findings, we repeat this analysis after removing all participants

who adopted an extreme strategy from the data (number of coins collected between 1 and 9). For the same reasons as above, we use Wald tests to assess if parameters associated with explanatory variables have a non-zero effect.

We include data from all eight experimental conditions in this analysis and do not consider interactions between explanatory variables, as we focus on the isolated effect of treatments, age and gender.

3. Results

We find clear evidence that collecting coins incurs a cost (in terms of evacuation success) in our experiment. The more coins participants collect, the smaller the proportion of participants that evacuate in time (Fig. 4). The observed negative correlation between collected coins and the fraction of escaped participants is stronger than expected by chance (Likelihood-ratio test on logistic regression, see Sect. 2.4; $X_1^2 = 667.07$, $p < 2 \times 10^{-16}$). Thus, these findings confirm that participants are taking a risk by collecting coins and justify our subsequent analysis that focusses in more detail on the number of coins collected by participants.
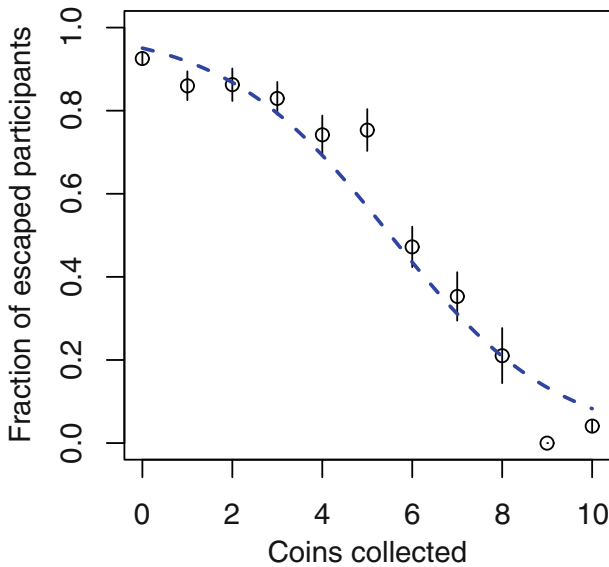

Figure 4. Coins collected plotted against the fraction of participants who reached the final exit within the time limit. Error bars show one standard deviation over 10,000 bootstrap samples of the data. The blue dashed line shows the fit of a logistic regression model (see Methods and Appendix A). The observed negative correlation between collected coins and the fraction of escaped participants is stronger than expected by chance (Likelihood-ratio test on logistic regression, see Sect. 2.4; $V_1^2 = 667.07$, $p < 2 \times 10^{-16}$).

Our data collection protocol means that participants are evenly spread across the eight experimental conditions (number of participants indicated in Fig. 5). The distributions of the number of coins collected in the different experimental conditions reveals that 44.6% of participants adopt an extreme strategy by either collecting all (18.1%) or none of the available coins (26.5%; Fig. 5). Considering the mean number of coins collected, experimental conditions that include treatment L (implied loss of coins, Fig. 5e–h) appear to result in higher values than in the remaining experimental conditions (Fig. 5a–d). In the following, we first investigate how our experimental treatments affect the prevalence of the two extreme strategies. Then, we test how the average level of risk-taking (mean number of
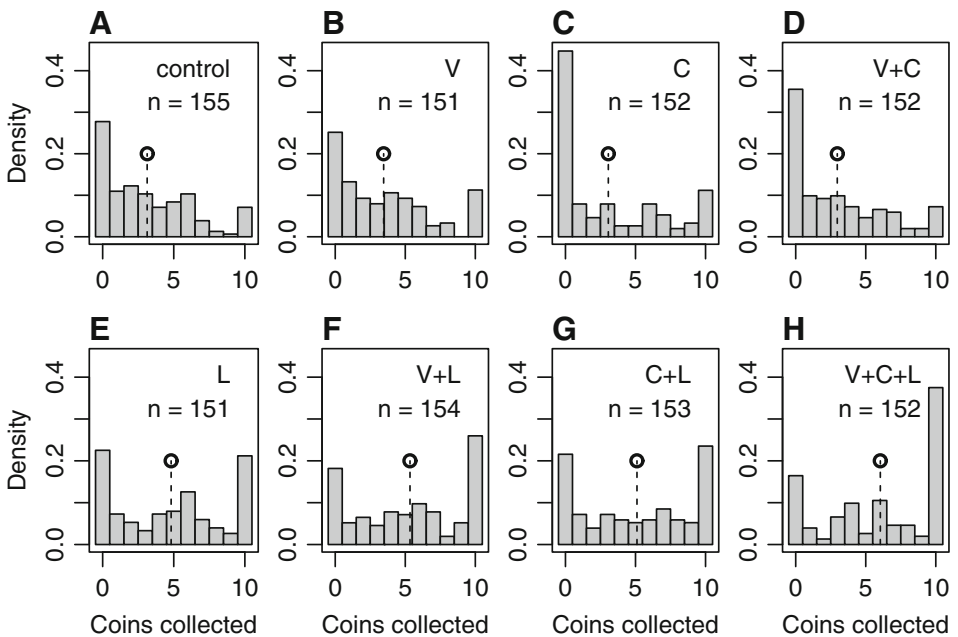

Figure 5. Distribution of the number of coins collected by participants for each experimental condition: (a) control condition; (b) final target not visible, treatment V; (c) crowd leaves late, treatment C; (d) combination of treatments V and C; (e) instructions imply coins were lost, treatment L; (f) combination of treatments V and L; (g) combination of treatments C and L; (h) combination of all experimental treatments V, C and L. Experimental treatments and number of participants included are indicated in the top right of each panel. Dashed vertical line segments highlighted with a circle indicate the mean number of coins collected for each experimental condition. For all experimental conditions, a high fraction of participants who collect no coins is clearly visible. In many experimental conditions there is an additional high prevalence of participants who collected all coins (panels E-H in particular).

collected coins) changes with experimental treatments and establish to what extent these effects can be explained by participants adopting the extreme strategies.

We find that participant age, as well as all treatments have significant effects on the fraction of participants who collect none of the available coins (all *p* values ≤ 0.05, see Table 1). Older participants are more likely not to collect any coins. Similarly, treatment C (crowd leaves late) leads to an increase in the fraction of participants who do not collect any coins. This is particularly visible in the absence of treatment L (compare Fig. 5a–c and b–d). One possible explanation for this result could be that when the crowd evacuates late (treatment C), there are initially no queues at the exits. This contrasts with the other experimental conditions when queues forming at the exits might deter participants from leaving immediately (see Fig. 3a, b). Rather than waiting in a queue, they may prefer to collect coins. Treatments V (lack of environment visibility) and L (implied loss) cause a decrease in the fraction of participants who collect no coins. While not very pronounced in Fig. 5 for treatment V, this trend is clearly visible for treatment L (compare Fig. 5a–d to e–h).

For the other extreme strategy in which participants collect all available coins, only treatments V (lack of environment visibility) and L (implied loss) have effects with *p* values ≤ 0.05 (Table 2). Both treatments lead to an increase in the fraction of participants adopting the strategy. The low and comparatively high effect sizes for treatments V and L, respectively (Table 2), suggest and Fig. 5 confirms, that this effect is not very clear for treatment V, but very pronounced for treatment L (compare Fig. 5a–d to e–h).

Having established that our experimental treatments affect the prevalence of extreme coin collection strategies, we next investigate the average level of risk participants take. Considering all data, including participants who adopt an extreme strategy, we find that only age and treatment L (implied loss) have significant

Table 1
Effect of Age, Gender and Experimental Treatments on the Proportion of Participants Who Collect No Coins

Explanatory factor	Parameter estimate ± s.e.	Wald test statistic	*p* value	Effect size against baseline
Intercept	− 1.53 ± 0.23	− 6.73	**1.65 × 10⁻¹¹**	NA
Age (years)	0.03 ± 0.01	4.71	**2.47 × 10⁻⁶**	0.07
Gender (male)	− 0.06 ± 0.14	− 0.38	0.70	− 0.01
Treatment V (visibility)	− 0.34 ± 0.14	− 2.35	**0.02**	− 0.07
Treatment C (crowd movement)	0.36 ± 0.15	2.47	**0.01**	0.08
Treatment L (implied loss)	− 0.74 ± 0.15	− 5.02	**5.12 × 10⁻⁷**	− 0.13

Results from logistic regression using the Boolean response variable indicating whether participants collect no coins (value 1) or not (see Appendix B). Effect sizes are shown against a baseline proportion of 0.30 of participants who collect no coins for a female aged 23 years (median age) with no experimental treatments present. The age effect size is shown for an age increase of 10 years. Analysis includes 1081 data points (for 139 participants age and/or gender information was missing). *p* values ≤ 0.05 are shown in bold

Table 2
Effect of Age, Gender and Experimental Treatments on the Proportion of Participants Who Collect All Coins

Explanatory factor	Parameter estimate ± s.e.	Wald test statistic	*p* value	Effect size against baseline
intercept	− 2.38 ± 0.29	− 8.26	$< 2.0 \times 10^{-16}$	NA
Age (years)	− 0.01 ± 0.01	− 1.78	0.07	− 0.01
Gender (male)	0.30 ± 0.17	1.82	0.07	0.02
Treatment V (visibility)	0.33 ± 0.17	1.99	**0.05**	0.02
Treatment C (crowd movement)	0.17 ± 0.16	1.01	0.31	0.01
Treatment L (implied loss)	1.36 ± 0.18	7.59	$\mathbf{3.25 \times 10^{-14}}$	0.14

Results from logistic regression using the Boolean response variable indicating whether participants collect all coins (value 1) or not (see Appendix B). Effect sizes are shown against a baseline proportion of 0.06 of participants who collect all coins for a female aged 23 years (median age) with no experimental treatments present. The age effect size is shown for an age increase of 10 years. Analysis includes 1081 data points (for 139 participants age and/or gender information was missing). *p* values ≤ 0.05 are shown in bold

effects on the average number of coins participants collect (both *p* values ≤ 0.05; Table 3). With increasing age, participants collect fewer coins, but treatment L leads to an average increase of about 2 coins that are collected (Table 3). The effects of treatments V and C on the average number of coins collected is insignificant. This suggests that for these treatments, the effects on the proportion of participants who adopt extreme strategies discussed above are not strong enough to result in significant changes in average levels of risk-taking.

Table 3
Effect of Age, Gender and Experimental Treatments on the Number of Coins Collected by Participants, Including Data on Participants Who Collected All or No Coins

Explanatory factor	Parameter estimate ± s.e.	Wald test statistic	*p* value	Effect size against baseline (coins)
intercept	1.33 ± 0.10	12.77	$< 2.0 \times 10^{-16}$	NA
Age (years)	− 0.01 ± 0.003	− 3.42	$\mathbf{6.28 \times 10^{-4}}$	− 0.29
Gender (male)	0.10 ± 0.06	1.51	0.13	0.30
Treatment V (visibility)	0.10 ± 0.06	1.51	0.13	0.30
Treatment C (crowd movement)	− 0.01 ± 0.06	− 0.17	0.87	− 0.03
Treatment L (implied loss)	0.53 ± 0.06	8.22	$\mathbf{< 2.0 \times 10^{-16}}$	2.05

Results from negative binomial regression using the number of coins collected by participants as the response variable (see Appendix C). Effect sizes are shown against a baseline of 2.97 coins for a female aged 23 years (median age) with no experimental treatments present. The age effect size is shown for an age increase of 10 years. Analysis includes 1081 data points (for 139 participants age and/or gender information was missing). *p* values ≤ 0.05 are shown in bold

To establish how influential the adoption of extreme strategies is on the observed trends in average risk-taking levels, we repeat our analysis after removing all data from participants who collect all or no coins (the extreme strategies). We find that in this data set age has no significant effect, but as for the full data set, treatment L (implied loss) leads to a significant effect by increasing the number of coins collected by about one coin, on average (Table 4). This suggests that age-related effects are mainly attributable to the adoption of extreme strategies, whereas the effect of treatment L is robust and occurs for extreme and intermediate strategies.

4. Discussion

In summary, we present the results of a virtual experiment with over 1200 human participants in which we explore how three key factors (reduced knowledge of building layout, crowd evacuation behaviour, and loss-aversion) influence the level of risk evacuees take by collecting objects (virtual 'coins') and thus causing a delay in the pre-movement phase of simulated evacuations. We find that many participants display extreme strategies of complete risk aversion (no coins collected, 26.5% of participants) or taking the highest risk possible (all coins collected, 18.1% of participants). While the movement of the simulated crowd (treatment C) and the information participants have about the location of exits (treatment V) affect the adoption of these extreme strategies, only the suggestion that participants can recover lost objects rather than gaining them (treatment L) has a robust effect on the mean level of risk-taking regardless of extreme strategies. Gender has no effect, but older participants are on average more risk averse.

Table 4
Effect of Age, Gender and Experimental Treatments on the Number of Coins Collected by Participants, Removing Data from Participants Who Collected All or No Coins

Explanatory factor	Parameter estimate ± s.e.	Wald test statistic	*p* value	Effect size against baseline (coins)
intercept	1.35 ± 0.07	19.77	**< 2.0 × 10⁻¹⁶**	NA
Age (years)	− 0.001 ± 0.002	− 0.57	0.57	− 0.05
Gender (male)	0.04 ± 0.04	0.87	0.38	0.15
Treatment V (visibility)	− 0.02 ± 0.04	− 0.36	0.72	− 0.06
Treatment C (crowd movement)	0.06 ± 0.04	1.49	0.14	0.25
Treatment L (implied loss)	0.22 ± 0.04	5.07	**4.02 × 10⁻⁷**	0.92

Results from negative binomial regression using the number of coins collected by participants as the response variable (see Appendix C). Effect sizes are shown against a baseline of 3.76 coins for a female aged 23 years (median age) with no experimental treatments present. The age effect size is shown for an age increase of 10 years. Analysis includes 610 data points (for 66 participants age and/or gender information was missing). *p* values ≤ 0.05 are shown in bold

However, this age-effect appears to be largely caused by differences in adopting extreme strategies.

It would be interesting to further investigate why participants adopt extreme strategies. It is possible that people who collected no coins take the evacuation context seriously, remember instructions on not collecting personal belongings in evacuations and therefore disregard the coins. The other extreme strategy might be caused by participants viewing the experiment as a game causing them to approach collecting all coins as a challenge. Asking participants about their motivation after completing the experiment could help to explain this finding. While we did not perform such a survey systematically, we anecdotally observed that some participants commented on wanting to win the coin collection challenge after the experiment.

Previous work suggests that there are gender differences in the level of risk-taking in pedestrians [30] and earlier virtual experiments also show gender effects but suggest these could be related to differences in exposure to computer games [23]. In contrast, we find no effect of gender on risk-taking in our experiment. The age effects we find are interesting, but we suggest that a better understanding of extreme strategy adoption is needed before interpreting them further. For example, participants' exposure to computer games could explain differences in strategy adoption. Frequent users of computer games may be more confident in their ability to collect all coins. Computer games may be played more frequently by younger participants and this could thus explain the age effects we observe. Similar factors could also explain the absence of a difference in risk-taking between genders. Since we recruited participants in a science museum, our participant pool may not be fully representative of society, although it is likely to be representative for museum visitors. We did not collect data on such confounding factors or on how frequently participants play computer games. As such, we suggest that while the relationship of age with risk-taking and the lack of a link between gender and risk taking we found are interesting and could be investigated further, they should only be considered directly within the context of our participant pool and virtual environment.

One approach to test or account for behavioural differences across individuals (e.g. related to gender, age or computer literacy) would be to ask participants to complete our experiment in multiple or even all experimental conditions. On the one hand, this should ensure that the distribution of individuals' characteristics is similar across experimental conditions which would facilitate assessing the effect of our experimental treatments, as well as differences in individual-specific responses to treatments. On the other hand, in this approach we would have to account for participants becoming habituated to the experiment which could itself depend on individuals' characteristics. We suggest that with over 150 participants per experimental condition, our data set is sufficient for the preliminary findings we present here.

Social influence occurs when individuals respond to the behaviour of others and it is an important factor that needs to be considered in research on pre-movement times in evacuations [10, 21]. In our experiment, altering the behaviour of the simulated crowd (treatment C) has no robust effect on the mean level of risk-taking.

This could have different causes. The scenario in our experiment means that there is no ambiguity on when the evacuation starts. This contrasts with other scenarios, where social influence mainly occurs when responding to others whilst interpreting an alarm [10] or deciding on where to go [21]. In addition, the simulated crowd only provides information on the evacuation timing. Social influence may affect the level of risk-taking more strongly if others provide direct examples of risk-taking behaviour that can be copied. Our experiment simulates a situation that may be unusual in real life, where only one person, the participant, can collect objects. People may be more likely to collect objects if they observe others doing it and vice-versa. One strength of our experimental paradigm is that it could easily be adapted to systematically test such additional scenarios. However, it is important to keep in mind that the realism of human social responses to simulated agent's behaviour is unclear, despite promising indications [31].

Even though previous work has identified familiarity with buildings as a determinant for pre-movement delays [5], in our experiment providing less information on the location of the final exit (treatment V) only affected the adoption of extreme strategies, but not the mean level of risk-taking. The layout of our virtual environment is simple and participants view it from a top-down perspective. In reality, a similar level of information would only be available if evacuees had familiarised themselves with the room layout by (for example) studying a floor plan. For these reasons, we suggest that our findings with regards to treatment V should be treated with care. Familiarity with buildings is likely to include cognitive processes that we cannot capture with our experimental setup [32].

The most robust effect on the level of risk-taking of participants in our experiment is caused by altering the quality of attachment to the objects that can be collected (treatment L). Motivated by prospect theory [28], we imply to participants via the instructions given to them that rather than being able to gain objects, they have lost objects they can recover. This extends previous work which suggests that the level of attachment to objects is crucial for determining the delays evacuees are willing to accept to collect them [18]. We find a loss-aversion effect for abstracted objects ('coins'). If this effect is not an artefact of our virtual experiment, it is certainly possible that it will be considerably stronger in reality for objects that evacuees are personally attached to. Current solutions to this problem aim to train (e.g. announcements on planes [20]) or to inform (e.g. wearable devices warn evacuees of delays [18]). It may not be possible to alter the quality of attachment people feel to objects in all cases, because of objects' emotional or other significance, but in some circumstances targeted interventions could be possible. For example, an additional measure could be to ensure that people in high-risk environments always carry with them important small items such as wallets, phones, passports and keys, therefore removing the urge to collect them and cause possible evacuation delays. In workplace environments, insurance and data backup policies could remove employees' need to collect laptop computers, data storage devices or work phones.

The level of attachment to objects, such as phones, wallets or house keys may differ across individuals. In our experiment, we deliberately include abstracted objects, virtual 'coins', that participants can collect, to avoid such differences

across individuals in attachment to more realistic items. While we intend to study the effect of treatment L against a less variable baseline with this approach, it also makes our experiment more abstract and may also contribute to participants treating the experiment as a game. For more life-like experiments, it would be advisable to consider more realistic objects that evacuees can collect.

The result that treatment L in our experiment leads to participants collecting more virtual coins and therefore to an increase in average pre-movement delays is an important proof of principle in its own right. This could have implications for how pre-movement delays are implemented in evacuation models. Typically, these delays are either modelled by approximations of empirically derived distributions or mechanistically by modelling individual-level decision-making [2, 14–17]. Based on our findings, we suggest that the global mean of pre-movement delay distributions in such models should routinely be varied over a range of values. In this way, the consequences of potential systematic effects based on occupant characteristics or targeted interventions could be assessed.

All controlled experimental studies on behaviour present abstractions of reality to participants. Ecological validity, or the extent to which findings from such studies extend to behaviour in the real world, is therefore a key issue. It is likely that our virtual environment experiment has lower ecological validity than high-fidelity unannounced evacuation drills undertaken in a real environment with a real crowd of people. Thus, our findings provide new insights into risk-taking behaviour in evacuations, but we cannot guarantee that they apply immediately to real-world contexts. Rather than employing them to directly study human behaviour in fire evacuation scenarios, we suggest virtual experiments can be used in different ways before their real-world validity is established. Our study demonstrates that virtual experiments make it possible to quickly, cheaply and safely conduct studies on human behaviour in crowd evacuations with large numbers of participants. Conducting experiments online, potentially using mobile devices, could make these cost and time savings even more dramatic [33] and our experiment can easily be extended to test additional scenarios. Based on this, we suggest a key use for virtual experiments lies in quickly and cheaply scoping behavioural responses in a broad range of scenarios to efficiently identify contexts that should be investigated further in more life-like experiments or evacuation drills. This is particularly relevant for scenarios that are difficult to study due to ethical or safety concerns (e.g. stressful evacuations [26] or ones under limited visibility conditions [1]).

The level of engagement provided by interactive virtual experiments, could also be used in education or health and safety training. Current education and training for evacuations is often limited to auditory, textual or video-based information (e.g. on planes [20]). Virtual environments would make it possible for people to interactively experience evacuation scenarios and to make mistakes they can learn from in a safe environment. Research on education techniques suggests that such gamification could lead to better engagement and improved learning [34].

In addition to exploring these uses of virtual environments for human behaviour in evacuations, future work could also investigate the decision-making process in more detail. For example, an interesting question is whether evacuees

decide on the number of objects to collect at the start of the evacuation, or if they adapt their decision dynamically based on new information received from the environment. Individuals may trade-off the utility of different actions [28] or use simple heuristics [35], such as 'leave when I have collected half of the objects', to decide when and how to evacuate. While our experiment focusses on individual-level decisions, many evacuation scenarios involve more than one person and reciprocal actions. For instance, in the context of collecting objects, individuals may only collect another item if someone else does the same or if they are competing with others. Helping behaviour in evacuations could also include reciprocal actions, such as people opening doors for each other [23]. Such reciprocal actions that are based on trading off perceived costs and benefits for outcomes could be approached using concepts from game theory and exploring the potential links of behaviour in evacuations to game theory could be an interesting avenue for future research.

5. Conclusions

Our study is one of the first to experimentally investigate pre-movement delays in evacuations. As our experiment is set within a virtual environment and because it presents a substantial abstraction of real-life situations, we suggest throughout that the findings from our work should not be over-interpreted and applied directly to real-world contexts. Therefore, we suggest our work and findings on the role of loss aversion, social influence, familiarity with buildings, age and gender on the level of risk people take by collecting objects before evacuating, act as a starting point with testable quantitative hypotheses that can be used to inform further studies. Despite these caveats, we suggest virtual environment experiments are a useful tool for studying human behaviour in evacuations. They facilitate a highly controlled, safe and cheap approach to research that is ideal for preliminary investigations, such as ours. Furthermore, it may be possible to adapt such virtual environment experiments for use in fire safety training, as they offer an engaging and interactive experience.

Acknowledgements

Open Access

Appendix A: Logistic Regression on Participants Evacuating on Time

Here, we describe the logistic regression model used to investigate the relationship between the number of collected coins and whether participants evacuated within the time limit.

Let W_i be a Boolean variable that indicates whether participant i evacuates within the time limit ($W_i = 1$) or not ($W_i = 0$) and let k_i denote the number of coins collected by this participant ($k_i \in [0, 10]$). In logistic regression, we assume W follows a binomial distribution, such that $W_i \sim Binom(1, \mu_i^A)$. The quantity μ_i^A is linked to the linear predictor, $\alpha_0 + \alpha_1 k_i$, where α_0 and α_1 are model parameters, via the inverse of the logit link function, such that: $\mu_i^A = \frac{1}{1+\exp(\alpha_0 + \alpha_1 k_i)}$. In effect, we model the probability for participants to evacuate within the time limit $P(W_i = 1) = \mu_i^A$. We fit this model to our data using a standard and well-established Maximum Likelihood approach (as implemented in the R programming environment [29]).

To assess if the number of coins collected has a non-zero effect on $P(W_i = 1)$, we perform a Likelihood ratio test on the null Hypothesis $H_0 : \alpha_1 = 0$ (see also Methods, Results and Fig. 3).

Appendix B: Logistic Regression on Extreme Strategy Adoption

Here, we describe the logistic regression model used to investigate the factors influencing participants' adoption of one of the extreme strategies in which they either collect all coins or no coins.

We measure the Boolean variable Y_i that indicates whether participant i collects no coins ($Y_i = 1$) or some coins ($Y_i = 0$; Table 1). We also repeat the following analysis using Y_i to measure whether participant i collects all coins or not (Table 2).

We assume Y follows a binomial distribution, $Y_i \sim Binom(1, \mu_i^B)$. The quantity μ_i^B is linked to the linear predictor, $X_i \beta$, of the model via the inverse of the logit link function, such that: $\mu_i^B = \frac{1}{1+\exp(X_i \beta)}$. Here β is a 6×1 vector of parameters and X_i is a 1×6 matrix containing the values of the explanatory variables, as evi-

dent from Eq. 1 below. The linear predictor of the model includes the additive effects of the explanatory variables we consider in our model:

$$X_i\beta = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \beta_3 x_{3,i} + \beta_4 x_{4,i} + \beta_5 x_{5,i} \tag{1}$$

Here, β_0 is the intercept and $\beta_1 \ldots \beta_5$ are the parameters capturing the effects of the different explanatory variables. Specifically, $x_{1,i}$ measures the age of participant i in years, $x_{2,i}$ is a dummy variable that takes value 1 if participant i is male and value 0 otherwise. Similarly, the remaining explanatory variables $x_{3,i}$ to $x_{5,i}$ are dummy variables indicating whether a treatment is present or not. In Tables 1 and 2, we report the estimates for $\beta_0 \ldots \beta_5$ obtained from Maximum Likelihood Estimation. As explained in the Methods, we use Wald tests to separately test for each parameter the null hypothesis that the parameter is equal to zero. Model fitting and statistical tests were performed using the default implementation in the R programming environment [29].

Appendix C: Negative Binomial Regression on Number of Coins Participants Collect

Here, we describe the negative binomial regression model used to investigate the factors influencing the number of coins participants collected in our experiment.

For this analysis, we use k_i, the number of coins collected by participant i, as the response variable. We perform negative binomial regression on all data ($k_i \in [0, 10]$, Table 3) and on data from which participants who adopt extreme strategies have been removed ($k_i \in [1, 9]$, Table 4).

We assume k_i follows a negative binomial distribution and use the log link function to compute the mean μ_i^C of this distribution from the linear predictor $X_i\gamma$, such that $\mu_i^C = \exp(X_i\gamma)$. We use the same linear predictor as in Appendix B, Eq. 1:

$$X_i\gamma = \gamma_0 + \gamma_1 x_{1,i} + \gamma_2 x_{2,i} + \gamma_3 x_{3,i} + \gamma_4 x_{4,i} + \gamma_5 x_{5,i} \tag{2}$$

The explanatory variables are same as in Appendix B and we report the estimates for the parameters $\gamma_0 \ldots \gamma_5$ obtained from Maximum Likelihood Estimation in Tables 3 and 4. We refer the reader to the detailed documentation of the R programming environment (function 'glm.nb' [29]) and to [36] for details on the model structure and fitting procedure for negative binomial regression.

References

1. Ronchi E, Nilsson D, Kojić S, Eriksson J, Lovreglio R, Modig H, Walter AL (2016) A virtual reality experiment on flashing lights at emergency exit portals for road tunnel evacuation. Fire Technol 52:623–647
2. Lovreglio R, Fonzone A, dell'Olio L (2016) A mixed logit model for predicting exit choice during building evacuations. Transp Res Part A 92:59–75

3. Zhang J, Seyfried A (2014) Quantification of bottleneck effects for different types of facilities. Transp Res Proc 2:51–59

4. Schadschneider A, Klingsch W, Kluepfel H, Kretz T, Rogsch C, Seyfried A (2009) Evacuation dynamics: empirical results, modeling and applications. In: Meyers RA (ed) Encyclopedia of complexity and systems science Springer, Heidelberg, pp 3142–3176

5. Proulx G (2002) Movement of people: the evacuation timing. In: DiNenno PJ (ed) SFPE handbook of fire protection engineering, 3rd edn. National Fire Protection Association, Quincy, pp 343–366

6. Zhao CM, Lo SM, Zhang SP, Liu M (2009) A post-fire survey on the pre-evacuation human behavior. Fire Technol 45:71–95

7. Kobes M, Helsloot I, De Vries B, Post JG (2010) Building safety and human behaviour in fire: a literature review. Fire Saf J 45:1–11

8. Johnson CW (2005) Lessons from the evacuation of the World Trade Center, Sept 11th 2001, for the future development of computer simulations. Cogn Technol Work 7:214–240

9. Shields TJ, Boyce KE (2000) A study of evacuation from large retail stores. Fire Saf J 35:25–49

10. Nilsson D, Johansson A (2009) Social influence during the initial phase of a fire evacuation—analysis of evacuation experiments in a cinema theatre. Fire Saf J 44:71–79

11. Liu M, Lo SM (2011) The quantitative investigation on people's pre-evacuation behavior under fire. Autom Constr 20:620–628

12. Lovreglio R, Ronchi E, Nilsson D (2016) An evacuation decision model based on risk perception, social influence and behavioural uncertainty. Simul Model Pract Theory 66:226–242

13. Purser DA, Bensilum M (2001) Quantification of behaviour for engineering design standards and escape time calculations. Saf Sci 38:157–182

14. Kuligowski E (2013) Predicting human behavior during fires. Fire Technol 49:101–120

15. Lovreglio R, Ronchi E, Nilsson D (2015) A model of the decision-making process during pre-evacuation. Fire Saf J 78:168–179

16. Reneke PA (2013) Evacuation decision model. US Department of Commerce, National Institute of Standards and Technology. http://dx.doi.org/10.6028/NIST.IR.7914

17. Gwynne SM, Hulse LM, Kinsey MJ (2016) Guidance for the model developer on representing human behavior in egress models. Fire Technol 52:775–800

18. D'Orazio M, Bernardini G (2014) An experimental study on the correlation between "attachment to belongings" "pre-movement" time. In: Weidmann U, Kirsch U, Schreckenberg M (eds) Pedestrian and evacuation dynamics 2012 Springer, Berlin, pp 167–178

19. D'Orazio M, Longhi S, Olivetti P, Bernardini G (2015) Design and experimental evaluation of an interactive system for pre-movement time reduction in case of fire. Autom Constr 52:16–28

20. IATA Cabin Operations Safety Task Force (2015) Cabin operations safety best practices guide, 2nd edn. International Air Transport Association. http://www.iata.org/publications/Documents/cabin-operations-safety-bp-guide-2015.pdf. Accessed 11 Apr 2017

21. Bode NWF, Holl S, Mehner W, Seyfried A (2015) Disentangling the impact of social groups on response times and movement dynamics in evacuations. PLoS ONE 10:e0121227

22. Kinateder M, Ronchi E, Nilsson D, Kobes M, Müller M, Pauli P, Mühlberger A (2014) Virtual reality for fire evacuation research. In: Ganzha M, Maciaszek L, Paprycki M (eds) 2014 Federated conference on computer science and information systems (FedCSIS). IEEE, pp 313–321

23. Bode NWF, Miller J, O'Gorman R, Codling EA (2015) Increased costs reduce reciprocal helping behaviour of humans in a virtual evacuation experiment. Sci Rep 5:15896
24. Kinateder M, Müller M, Jost M, Mühlberger A, Pauli P (2014) Social influence in a virtual tunnel fire–influence of conflicting information on evacuation behavior. Appl Ergon 45:1649–1659
25. Kobes M, Helsloot I, de Vries B, Post J (2010) Exit choice, (pre-) movement time and (pre-) evacuation behaviour in hotel fire evacuation—behavioural analysis and validation of the use of serious gaming in experimental research. Proc Eng 3:37–51
26. Bode NWF, Wagoum AUK, Codling EA (2015) Information use by humans during dynamic route choice in virtual crowd evacuations. R Soc Open Sci 2:140410
27. Bode NWF, Wagoum AUK, Codling EA (2014) Human responses to multiple sources of directional information in virtual crowd evacuations. J R Soc Interface 11:20130904
28. Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. J Risk Uncertain 5:297–323
29. R Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. Accessed 27 Jan 2017
30. Granié MA (2009) Effects of gender, sex-stereotype conformity, age and internalization on risk-taking among adolescent pedestrians. Saf Sci 47:1277–1283
31. Drury J, Cocking C, Reicher S, Burton A, Schofield D, Hardwick A, Graham D, Langston P (2009) Cooperation versus competition in a mass emergency evacuation: a new laboratory simulation and a new theoretical model. Behav Res Methods 41:957–970
32. Andresen E, Haensel D, Chraibi M, Seyfried A (2016) Wayfinding and cognitive maps for pedestrian models. In: Knoop VL, Daamen W (eds) Traffic and granular flow'15 Springer, Berlin, pp 249–256
33. Brown HR, Zeidman P, Smittenaar P, Adams RA, McNab F, Rutledge RB, Dolan RJ (2014) Crowdsourcing for cognitive science—the utility of smartphones. PLoS ONE 9:e100662
34. Howard-Jones PA, Jay T, Mason A, Jones H (2016) Gamification of learning deactivates the default mode network. Front Psychol 6:1891
35. Seitz MJ, Bode NWF, Köster G (2016) How cognitive heuristics can explain social interactions in spatial movement. J R Soc Interface 13:20160439
36. Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer, New York. ISBN 0-387-95457-0