Please cite as:
Chittaro L., Corbett C.L., McLean G.A., Zangrando N., Safety knowledge transfer through mobile virtual reality: A study of aviation life preserver donning, *Safety Science*, 102, February 2018, pp. 159-168.

Safety Knowledge Transfer through Mobile Virtual Reality: a Study of Aviation Life Preserver Donning

Luca Chittaro[1], Cynthia L. Corbett[2], G.A. McLean[2], Nicola Zangrando[1]

1. Human-Computer Interaction Lab (HCI lab), University of Udine, Italy
2. Prime Stratagem, LLC

Abstract

Aviation safety knowledge is a key factor in determining how passengers will respond in an emergency, but the effectiveness of the tools (preflight safety briefing, safety briefing card) used by airlines to educate passengers about safety has been shown to be lacking. This paper explores how one of these tools could be made interactive in order to increase its effectiveness. In particular, we use Virtual Reality (VR) techniques, adapting them to the constraints imposed by on-board aircraft use, such as usage on non-immersive, small displays. As a practical application, the paper examines aviation life preserver donning, which the literature has shown to be particularly difficult for passengers. To evaluate the proposed mobile VR tool, we contrasted it with the traditional safety briefing card in a between-groups study with 68 participants, age 20-24, focusing on different aspects of effectiveness. The results of the study show that the participants who used the mobile VR tool were able to transfer the presented safety knowledge to the real world, and don an aviation life preserver significantly faster and with fewer errors than participants who used the traditional briefing card. Moreover, these objective results were consistent with subjective ratings by participants; the mobile VR tool was perceived as significantly more engaging, simpler, and more effective than the traditional briefing card. Finally, participants who used the mobile VR tool attained a higher level of self-efficacy. The generalizability of these results would benefit with additional work aimed at an older age cohort that would ostensibly be less familiar with interactive VR technology.

Keywords: aviation safety, virtual reality, mobile devices, safety education, safety training.

1. Introduction

Aviation safety knowledge is a key factor in determining how passengers will respond in an emergency (Chang & Liao, 2009; Muir & Thomas, 2004; Thomas, 2003; Edwards, 1990), greater knowledge making them better able to handle the situation and better prepared to utilize the emergency equipment in the cabin. The two knowledge-based tools routinely used by airlines to educate passengers about safety are the preflight safety briefing and the safety briefing card. Unfortunately, both of them have been shown to suffer from a serious lack of effectiveness, as shown by empirical studies conducted with passengers (Corbett & McLean, 2007; Corbett, McLean, & Cosper, 2008; Seneviratne and Molesworth, 2015), interviews of aircraft accident survivors (NTSB, 2000; Chang & Yang, 2011), and accident reports (e.g., NTSB, 2010).

One reason for the lack of effectiveness of current tools is that they are not engaging enough, and as a result only a minority of passengers pays attention to them. Corbett and McLean (2007) found that only 30% to 40% of people reported that they attended to safety information, and the NTSB (2000, 2010) found similar results in accident investigations. The other major reason is lack of comprehension, even among those passengers who reportedly paid attention to the briefings and engaged the safety briefing card (Corbett & McLean, 2007; Corbett, McLean, & Cosper, 2008). The ineffectiveness of these current tools has been confirmed by the experiences of aircraft accidents survivors. For example, in the interviews of 110 accident survivors conducted by Chang and Yang (2011), only 14% found the preflight safety briefing to be useful for successfully surviving the accident, and only 16% found the safety briefing card to be effective. The majority of survivors said that they did not believe they were adequately instructed.

In an attempt to face these shortcomings, some airlines have begun to employ preflight safety briefing videos that aim at being more engaging by using humour or employing celebrities as speakers, although the effectiveness of these efforts also remains in doubt. Seneviratne and Molesworth (2015) have provided an initial evaluation of these approaches, contrasting the traditional safety briefing video with these new types of safety videos, one based on humour and the other presented by a celebrity. Although the humorous video did slightly better, the overall results obtained by the three types of videos were considered “alarming” (participants recalled only about half of the safety information), leading the authors to conclude that “airlines and aviation authorities need to rethink the way in which they convey safety critical information to passengers.”

Other researchers (Muir & Thomas, 2004, Chang & Liao, 2009) and reports from aviation safety authorities (Cosper & McLean, 2004; NTSB, 2010) have made similar recommendations, recommending instead more creative approaches, such as interactive technologies (Cosper & McLean, 2004; Chittaro, 2017) or hands-on safety education exhibits at airports (Chang & Yang, 2011).

1.1 Exploring Mobile Virtual Reality for Aviation Safety Briefings

The goal of this paper is to explore and evaluate the possible effectiveness of a novel Virtual Reality (VR) approach to make safety briefings interactive. Among the interactive technologies one could consider, VR appears as a promising choice, based on its increasingly important role in safety training spanning several domains, including fire safety (Cha et al., 2012; Smith & Ericson, 2009), mining industry operations (Grabowski & Jankowski, 2015), construction workers safety (Guo et al., 2012), naval safety (Stone, Caird-Daley & Bessell, 2009), road safety (Li & Tai, 2014), disaster preparedness (Andreatta et al., 2010), and emergency medical response (Cohen et al., 2013). In addition to showing the effectiveness of VR as a safety training tool, the literature has also emphasized the engagement that VR can create in its users, considering it as a factor that can improve safety training (Grabowski & Jankowski, 2015).

In consideration of aircraft passenger safety briefings, however, two major barriers must be overcome in order to use VR for creating novel types of safety briefings that passengers could use on-board the aircraft. The first concerns the equipment needed to use the interactive tool. Current VR safety training tools are often based on immersive hardware (such as head-mounted displays or multi-screen projections), and even tools that are not based on such special hardware require a personal computer with a good 3D graphics board. This makes it essentially impossible to offer current VR training tools to the passenger on-board an aircraft, restricting their usage to special ground facilities or, in the best case, to home and office environments. An interactive safety briefing tool for aircraft passengers should instead be designed for small screens and less powerful systems that are currently used in the aircraft cabin. The In-Flight Entertainment Systems (IFEs) mounted on the seats of some aircraft support interactive applications and could provide a first opportunity to offer interactive safety briefings. Unfortunately, such IFEs are currently available only on selected long-haul flights. Another opportunity is to exploit the widespread Personal Electronic Devices (PEDs), such as smartphones and tablets, that most passengers bring on-board. Regulatory constraints on PED usage have eased, as the latest FAA and EASA policy is to allow PED use by passengers during all phases of flight. Moreover, it opens up new opportunities for delivering safety knowledge, e.g. the interactive content could be sent together with electronic tickets or boarding passes that passengers already receive on their smartphones, making it possible to use it more discreetly than IFEs and before boarding the aircraft. To overcome the first barrier, this paper focuses on creating a VR tool that can run on the small touchscreens of mobile systems such as those of IFEs and PEDs, including common smartphones. For this reason, we use the term "Mobile VR" in the paper.

The second barrier concerns the design of the content of VR safety training tools, which are typically based on realistic simulations of the emergencies one should prepare for. In aviation safety, such an approach would likely result in 3D reconstructions of serious aircraft accidents and their effects on passengers. All previous work on using VR for safety training of aircraft passengers has followed this simulation approach (Chittaro & Buttussi, 2015; Chittaro, 2016). Unfortunately, although realistic accident simulation is effective when used in on-ground training, such fearful content is emotionally inappropriate for on-board use by passengers (Chittaro, 2016).

To overcome the second barrier, the design we evaluate in this paper aims at creating a reassuring experience that does not expose the user to fearful content as the previously mentioned simulations do. In particular, we focus on creating an interactive version of the non-interactive illustrations that are currently provided by airlines in safety briefing cards and videos.

1.2 Evaluating the Approach

As a practical application of the Mobile VR approach, we considered aviation life preserver donning, because it appears particularly difficult for passengers to understand. Indeed, recent studies have shown that illustrations used by airlines to present life preserver donning are difficult to comprehend, even when study participants are given an unlimited amount of time to study them (Corbett & McLean, 2008;

Weed, Corbett, & McLean, 2013). Moreover, the US Airways Flight 1549 accident, in which the aircraft was forced to ditch in the Hudson River after a bird strike, brought significant, fresh attention to the fact that passengers are not knowledgeable about life preservers and not well prepared for using them. In its accident investigation report, the NTSB (2010) noted that many passengers did not even retrieve the life preservers and, of those who did, the majority indicated that they had difficulty donning them. Errors in using life preservers can cause passenger deaths in conditions that would otherwise be survivable, as Chang and Liao (2010) exemplify in their description of two different aviation accidents.

To evaluate thoroughly our Mobile VR tool, this paper contrasts it with the traditional non-interactive illustrations used by airlines on safety briefing cards, and focuses on measuring different aspects of effectiveness. First, we measured knowledge transfer, because any instructional technique (traditional or computer-based) would be of limited value if people could not effectively apply the acquired knowledge to the real world (Carpenter, 2012). The word “transfer” indicates such application of knowledge. As Bertram, Moskaliuk & Cress (2015) recently pointed out in searching the training transfer literature, a lack of studies on the transfer of VR training to reality is apparent. Moreover, the few available studies are based on traditional VR set-ups (immersive or desktop) and, to the best of our knowledge, no study has been conducted on training transfer from Mobile VR. Second, we measured subjective perceptions of users in terms of simplicity and efficacy of the received safety instructions and level of engagement. Third, we included a measure of attitude change. As pointed out by Chang & Liao (2009), in addition to providing passengers with accurate cabin safety knowledge, aviation safety education must also cultivate positive passenger attitudes that could enhance their behavior in an emergency. An important positive attitude is self-efficacy, which can be defined as the confidence in one’s ability to perform a behavior. According to Social Cognitive Theory (Bandura 1997; Bandura, 2001), this belief significantly determines performance outcomes, and different people with similar skills may perform differently depending on differences in their conviction that they can successfully execute a required behavior. In particular, positive associations between safety training, self-efficacy and attitudes toward safety have been described in the literature (see Grau, Martínez, Agut, & Salanova, 2002; Katz-Navon, Naveh, & Stern, 2007, for summaries). Increasing self-efficacy is particularly important in aviation safety education, because passenger attitudes about aircraft accidents tend to be pessimistic and fatalistic; they believe that there is little hope of survival and/or shift the responsibility and capability of their safety to the cabin crew (Muir & Thomas, 2004). Actually, the majority of aircraft accidents is survivable, as shown by surveys of commercial jet airplane accidents (Cherry, 2013). Moreover, workload and time constraints in aircraft evacuations make it impossible for the crew to provide individual assistance to every passenger.

In the following, we outline the principles that guided the design of the Mobile VR tool (Section 2) and then we illustrate the tool in detail (Section 3). The study of effectiveness and its results are presented in Sections 4 and 5, respectively, while Section 6 provides conclusions and suggestions for future work.

2. General design approach

We organized the design of the proposed Mobile VR tool by identifying the following principles and using them as guidance:

Constructivism. In pedagogy, constructivists underline the role of a direct experience of the world as a fundamental factor in learning (Chittaro & Ranon, 2007). As pointed out by Harper, Hedberg, & Wright (2000), apart from reality, the most appropriate way to generate a context based on authentic learner activity can be through the use of interactive VR worlds within which the user can act.

Embodied action. Actions in the VR world should be performed by a virtual human character that clearly reflects how such actions are physically carried out in real-world settings. The interaction between the user and the character should be made as close and direct as possible; we employ no menus or indirect interaction techniques, but a technique in which the user touches the virtual object s/he wants to act on, and then moves the character’s hands to act on the object as if s/he were guiding the hands of a real human.

Immediate feedback and assessment. The interactive experience should allow the user to actively explore the situation depicted by the VR world, performing right and wrong actions. Actions should be immediately assessed by the application and the user receive instant feedback about his/her behavioral outcomes. In this way, simulation of actions enables the user to observe immediately the link between cause and effect (Fogg, 2003). Where needed, the user should also be able to ask for contextual tips about what action to perform. Since feedback and tips depend on the actions taken and the possible help asked for by the individual user, they make the virtual experience different for each user and support individualized learning.

Fidelity. Fidelity to real equipment and environment in VR training is important for knowledge transfer (Dorsey et al, 2009). For example, the application considered in this paper carefully reproduces the life preserver, the cabin environment, the movements of the character for donning the life preserver, and life preserver responses to manipulation.

Game-like design. The app should exploit digital game design techniques, which contribute to attract user attention and make the learning experience more informal and appealing. As pointed out by Zyda (2005), games with a serious purpose (often called "serious games") can be an effective tool to further training and education objectives. Repetitive rehearsal of the training procedures could be encouraged through a scoring system that invites the user to improve his/her performance. Serious game design should also focus on evoking user emotions, to make the learning experience more engaging and more memorable. For example, we tried to introduce some humorous elements in the interactive experience (the character can make funny remarks or movements in response to a user's right or wrong choice).

3. Application to aviation life preserver donning

We followed the previously described general principles in the creation of a Mobile VR application aimed at teaching how to don a life preserver (for conciseness, we will refer to it as "app" in the following). We developed the app using the Unity 4.5 game engine and the C# programming language. This section describes the app and how it is used in detail.

The app reproduces a full 3D aircraft cabin environment in which a user sees his/her character (a virtual passenger) in third-person view (Figure 1). The goal for the user is to make the character don the life preserver properly. Toward this purpose, the user quickly and easily controls the character by touching the screen with his/her finger and gestures, as follows:

- **Choosing an object**. To act on an object in the 3D world, the user first touches the object on the screen. For example, by touching the compartment under the seat (see Figure 1), the view automatically zooms in on the object and the hand of the character approaches the object for possible action (Figure 3a).

- **Acting on an object**. After choosing an object, moving the finger on the screen controls the hand (or hands) of the character that has (have) approached the object. For example, by keeping the finger on the screen and dragging away from the compartment, the user opens the compartment and retrieves the life preserver (Figure 3b).
- **Zooming out**. A double touch on the screen zooms the view out from the object, bringing it back to the level of the entire cabin environment.
- **Changing viewing angle**. When the character is not acting on an object, the user can optionally change the viewing angle through which s/he looks at the object or environment, by touching the screen and dragging in the desired direction. To help the user in managing the viewpoint, possible viewing angles are constrained to keep the view meaningful; the app sets the initial viewpoint so that the character and the object on which it is going to act are in front of the virtual camera (see Figure 1 for the initial viewpoint).The user can then change the viewpoint within a range of [-30,+30] degrees horizontally and [-12,+18] degrees vertically.

When s/he uses the app for the first time, the user is quickly introduced to the four interaction commands above via a brief interactive tutorial in which they see the character seated in the aircraft with a laptop placed on the seat on its side (Figure 2a). The tutorial first asks the user to try experimenting with changing the viewing angle (Figure 2a), then it asks him/her to focus on the laptop (Figure 2b), to lift and lower the screen of the laptop (Figure 2c), and finally, to zoom back to the full cabin environment (Figure 2d). Once the user has successfully completed the actions in the brief tutorial, the app starts the life preserver instructional phase, revealing that the goal is to don the life preserver (Figure 1).

Properly donning the life preserver, in the real-world as well as in the mobile 3D world, requires the user to perform the following sequence of actions which is illustrated in safety briefing cards:

1) *Open Compartment*: locate and open the life preserver compartment under the seat (Figure 3a),

2) *Pull out Pouch*: pull the life preserver pouch from the compartment (Figure 3b),

3) *Open Pouch*: open the pouch by pulling its tab (Figure 3c),

4) *Pull out/Unfold Vest*: pull the folded life preserver from the pouch and unfold it (Figure 3d),

5) *Slip Vest over Head*: slip the life preserver over the head in such a way that the connected waist strap and buckle hang in front of the body (Figure 3e),

6) *Pass around Strap*: locate the strap and pass it around the waist (Figure 3f),

7) *Buckle Strap*: buckle the strap (Figure 3g),

8) *Tighten Strap*: tighten the strap by pulling the yellow tab (Figure 3h).

The character gives negative feedback to the user whenever s/he touches a part of the life preserver or the character that is unrelated to the currently required action (e.g., touching the strap before slipping the life preserver over the head, touching the yellow tab when the strap is not buckled). To give negative feedback, the character shakes its head and makes a brief remark, randomly chosen among a set of pre-recorded voice files (“Nope!”, “Not now”, “Naaaaa”, “I don't think so”, “Not like this”, “Think better”, “Seriously?”, “Are you sure?”, “You're kidding me”, “No, try again”, “Wrong action”, “No way”, “This is not going to help”). This design choice was taken to make interaction with the character more realistic, believable, and closer to the behavior of real humans. Interaction with a character that is able to utter only a single remark would be perceived as repetitive, artificial, and not engaging.

If the user remains inactive without touching the screen for 6 seconds, a light bulb icon appears in the upper right corner of the screen (Figure 4a) with a “ding” sound to indicate that the app can offer a tip about how to proceed. If the user touches the icon, the tip appears in the upper part of the screen (Figure 4b).

When the user succeeds in completing the life preserver donning procedure, the virtual passenger first looks at itself with a satisfied facial expression and then congratulates the user for the achievement by giving a thumbs up with both hands (Figure 5) and makes an exclamation, randomly chosen among a set of pre-recorded voice files (“Awesome”, “Great job”, “Well done”, “Brilliant”).

Finally, the app shows the time taken (in seconds) to make the character don the life preserver. If the user tries the procedure multiple times, the app keeps track of his/her best time, indicating after each try if the new time is better or worse than the previous best time. This is also an implicit incentive to retry the procedure to see if s/he can improve.

4. Evaluation

We carried out a between-groups study, in which half of the participants (App Group) used the app as the instructional medium, while the other half (Card Group) used an airline safety briefing card that presented the donning procedure. Our hypotheses were: (i) the app supports knowledge transfer to the real world, doing it more effectively than the safety briefing card, and leading people who use the app to perform life preserver donning better in the real world, (ii) the app is more engaging for users than the card, (iii) users of the app will be more confident about donning a life preserver, and (iv) users will perceive the safety instructions received from the app as simpler and more effective than the same instructions presented by the card.

4.1 Materials

The app was run on a LG Nexus 5 smartphone (screen size: 4.95 inches, resolution: 1920x1080 pixels). The safety briefing card was A4-sized, printed in color. The pictorials included on the card referred to the same life preserver depicted in the app, and were taken from the safety briefing card currently employed by one of the largest world airlines.

For the real-world life preserver donning test (described in the Procedure section), we used a life preserver of the same type depicted in the two instructional media.

4.2 Participants

The evaluation involved a sample of 68 undergraduate, computer science students (61 males, 7 females) with no background on safety topics. Their age ranged from 20 to 24 (M=21.21, SD=0.64). None of the participants had ever donned or tried to don an aviation life preserver.

We assessed individual frequency of air travel by asking participants to count their number of flights in the last two years, as in (Corbett et al., 2008). We made it clear that each flight had to be counted individually; thus, a round trip from airport A to airport C via a connection through airport B comprises four flights. Flight frequency ranged from 0 to 10 (M=1.68, SD=2.54).

We also assessed participant self-efficacy with respect to donning aviation life preservers. For this we designed an 8-item questionnaire, (i) adapting items from well-known self-efficacy questionnaires such as the General Self-Efficacy (GSE) scale (Schwarzer & Jerusalem, 1995) to illustrate beliefs about using the life preserver ("I feel confident about my ability to don it", "I would be able to don it correctly", "I would be able to don it fast", "I believe I would be able to help other passengers in donning it"), and (ii) following the recommendation about self-efficacy assessment of (Luszczynska & Schwarzer, 2005) i.e., "I feel able to don the life preserver in time if the aircraft lands on water", "I would be able to carry out all actions needed to wear the life preserver… even if the situation puts me under heavy pressure", "…even in a serious emergency on a sinking aircraft", "…even if most passengers are screaming or crying". Each item was rated by participants on a 7-point Likert scale (1=not at all, 7=very). Answers were averaged and the reliability of the scale was confirmed by Cronbach's test (alpha=.91). Participant's self-efficacy before trying the instructional media ranged from 1.12 to 5.25 (M=3.72, SD=1.18).

Participants were assigned to the two instructional conditions in such a way that: (i) the proportion of men and women was almost identical (31 males and 3 females in the Card Group; 30 males and 4 females in the App Group), and (ii) the two groups were very similar in terms of age, frequency of air travel, and self-efficacy. Independent samples t-tests confirmed the lack of significant differences between the two groups for the three demographic variables.

4.3 Measures

4.3.1 Engagement

To measure the level of engagement experienced by participants, we administered a questionnaire that asked them to think about the instructional experience they had just received and rate their level of agreement on a 7-point Likert scale (1=not at all, 7=very) about five statements, worded in such a way that they could apply to both the app and the card. The five items were: "It was boring", "It was engaging", "It was fun", "The depicted situation looked real", "I felt immersed in the depicted situation”. After inverting the scale of the first item, the five ratings were averaged to form the “engagement” measure, whose reliability was confirmed by Cronbach’s test (alpha=.91).

4.3.2 Instructions simplicity

We measured the perceived simplicity of the instructions received by the instructional media by having respondents rate their level of agreement with three items on a 7-point Likert scale (1=not at all, 7=very). The items made the following statements about the instructions provided by the instructional media: “They are simple”, “They are easy to learn”, “They are easy to carry out”. Answers to the three items were averaged to form the “simplicity” measure, whose reliability was confirmed by Cronbach’s test (alpha=.83).

4.3.3 Instructions efficacy

We measured the perceived efficacy of the instructions received by the instructional media by having respondents rate their level of agreement with three items on a 7-point Likert scale (1=not at all, 7=very). The items made the following statements about the instructions provided by the instructional media: “They are useful for my safety”, “They are effective to face a water landing”, “They allow one to greatly reduce the probability of getting hurt in a water landing”. Answers to the three items were averaged to form the “efficacy” measure, whose reliability was confirmed by Cronbach’s test (alpha=.82).

4.3.4 Real-world performance: time and errors

To measure participants’ performance in donning a real aviation life preserver after using the card or the app, we asked them to perform the procedure taught by the instructional media. To carry out this test, we previously attached a compartment under the participant’s seat (Figure 6). A real aviation life preserver, sealed in its pouch (Figure 7), was stored in the compartment. We measured the total time it took participants to don the life preserver, including pouch retrieval and opening. To keep track of errors, we observed participants while they donned the life preserver, and checked on a computer in which of the eight steps (described in Section 3) of the procedure they made the error. For each step, we considered as an error any action that was unuseful to carry out the step. For steps 6, 7, and 8, omission of the step was also a possible error, while it was impossible to omit any of the previous five steps, because completion of each of them was necessary to proceed to the next one. A detailed description of the errors made by participants is provided in Table 1.

4.3.5 Self-efficacy

In addition to measuring self-efficacy before the test (pre-instruction), we re-tested self-efficacy after participants studied the instructional media (post-instruction) and after they donned the life preserver in the real world (post-donning), to measure possible changes in self-efficacy as a result of these activities. For all three measurements, we used the questionnaire described in Section 4.2.

4.4 Procedure

We told participants that the goal of the experiment was to evaluate instructional media that illustrate how to don an aviation life preserver. Participants completed the initial questionnaire (age, frequency of air travel, self-efficacy), then received the instructional media (card or app) assigned to them. We handed the instructional media to the seated participants, who held it until instructed to proceed.

Participants in the App Group received an explanation about how to interact with the touchscreen from the app itself through the brief interactive tutorial described in Section 3.

After studying the instructional media assigned to them, participants answered the questions about self-efficacy, engagement, perception of instruction simplicity and instruction efficacy. Then, we informed them that their seat was equipped with a real aviation life preserver (Figure 6), without telling them where it was specifically, and asked them to try to don it by putting into practice what they had learned from the instructional media. We told them to start the task only after receiving a verbal “go” command from the experimenter and to raise their hands when they felt they had completed the task. We gave the “go” command while simultaneously starting a chronograph, and stopped the chronograph when participants raised their hands. The life preserver was of the same type depicted by the instructional media. Finally, participants again completed the self-efficacy questionnaire, and were thanked for their participation.

5. Results

5.1 Engagement

Differences in engagement (Figure 8) were analyzed with a between-subjects ANOVA. The difference between the two groups was statistically significant, $F(1, 66)=79.66$, $p<0.001$, $\eta_p^2=0.47$. The app ($M=5.12$, $SD=1.10$) was more engaging than the card ($M=2.96$, $SD=1.24$).

5.2 Instructions simplicity

Differences in perceived simplicity of the instructions (Figure 9) were analyzed with a between-subjects ANOVA. The difference between the two groups was statistically significant, $F(1, 66)=15.00$, $p<0.001$, $\eta_p^2=0.19$. Instructions given by the app ($M=5.58$, $SD=1.16$) were perceived as simpler than those of the card ($M=4.40$, $SD=1.34$).

5.3 Instructions efficacy

Differences in perceived instructions efficacy (Figure 9) were analyzed with a between-subjects ANOVA. The difference between the two groups was statistically significant, $F(1, 66)=7.17$, $p<0.01$, $\eta_p^2=0.10$. Instructions given by the app ($M=4.97$, $SD=1.37$) were perceived as more effective than those of the card ($M=4.13$, $SD=1.22$).

5.4 Real-World Performance: Time

Differences in time required to don the real life preserver (Figure 10) were analyzed with a between-subjects ANOVA. Participants in the App Group ($M=35.22$ s, $SD=11.06$) were able to don the life preserver in less time than those in the Card Group ($M=42.81$ s, $SD=14.22$), the difference being statistically significant, $F(1, 66)=6.05$, $p=0.017$, $\eta_p^2=0.08$..

5.5 Real-World Performance: Errors

Table 1 reports, for each step of the donning procedure, how many users made an error at that step. The last line of the table reports the total number of errors made in each group. In four of the eight steps

(Open Compartment, Pull out Pouch, Pull out/Unfold Vest, Buckle Strap), no participants made errors. For the other four steps (Open Pouch, Slip Vest over Head, Pass around Strap, Tighten Strap), the table includes a detailed description of the errors made. For the four steps in which participants made errors, the differences between the two groups were analysed with the Mann-Whitney test. The difference in Open Pouch ($Z=-2.05$, $p<0.05$), Slip Vest over Head ($Z=-2.11$, $p<0.05$) and Total Errors ($Z=-1.99$, $p<0.05$) were statistically significant, with the App Group always making fewer errors than the Card Group.

5.6 Self-efficacy

Self-efficacy scores were analyzed by a 2 x 3 mixed design ANOVA, in which Group (App, Card) served as the between-subjects variable and Self-efficacy measurement time (pre-instruction, post-instruction, post-donning) served as the within-subjects variable. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2)=10.17$, $p=0.006$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon=0.87$). There was a statistically significant main effect of Self-efficacy measurement time, $F(1.75, 115.30)=50.29$, $p<0.001$, $\eta_p^2=0.43$, and a Group by Self-efficacy measurement time interaction effect, $F(1.75, 115.30)=3.34$, $p<0.05$, $\eta_p^2=0.05$. Following Cohen (2013), we investigated the interaction effect by testing the simple main effect of Self-efficacy measurement time for each group, employing a one-way repeated measures ANOVA, followed by pairwise comparisons using the Bonferroni test.

In the Card Group (Figure 11), mean self-efficacy was 3.67 (SD=1.17) pre-instruction, 4.64 (SD=1.16) post-instruction, and 4.73 (SD=1.72) post-donning. Repeated measures ANOVA with Greenhouse-Geisser correction (Mauchly's test $\chi^2(2)=13.82$, $p=0.01$, $\varepsilon=0.74$) revealed a statistically significant difference, $F(1.48, 48.86)=12.00$, $p<0.001$, $\eta_p^2=0.27$. Pairwise comparisons with the Bonferroni test revealed that the difference between the first measurement of self-efficacy (pre-instruction) and its two subsequent measurements was significant ($p<0.001$ and $p<0.01$, respectively), while there was no statistically significant difference between self-efficacy post-instruction and post-donning.

In the App Group (Figure 11), mean self-efficacy was 3.72 (SD=1.18) pre-instruction, 4.88 (SD=1.18) post-instruction, and 5.51 (SD=0.83) post-donning. Repeated measures ANOVA revealed a statistically significant difference, $F(2, 66)=54.98$, $p<0.001$, $\eta_p^2=0.63$. Pairwise comparisons with the Bonferroni test revealed that the difference between the first measurement of self-efficacy (pre-instruction) and its two subsequent measurements was significant ($p<0.001$ in both cases). Unlike the Card Group, the App Group showed a statistically significant increase in self-efficacy in the comparison of self-efficacy post-instruction and post-donning ($p=0.001$).

To compare the effects of group at each level of Self-efficacy measurement time, we performed a between-subjects ANOVA for each of its three levels. This revealed a statistically significant difference only at post-donning time, $F(1, 66)=5.70$, $p=0.02$, $\eta_p^2=0.08$, with higher self-efficacy in the App Group (M=5.51, SD=0.83) than in the Card Group (M=4.73, SD=1.72).

6. Discussion and Conclusions

The results of the study confirmed the hypotheses. Participants who had used the Mobile VR tool were able to transfer the safety knowledge to the real world, retrieving and donning the life preserver. Moreover, they were able to do it significantly faster and with fewer errors than participants who had used the traditional safety briefing card. This result is particularly important, because donning the life preserver efficiently is a fundamental survival factor in real emergencies (Corbett, Weed, Ruppel, Larcher & McLean, 2014). Moreover, these results were consistent with subjective ratings by participants; the Mobile VR tool was perceived as significantly more engaging, simpler and more effective than the traditional briefing card. To the best of our knowledge, our work is the first to propose and evaluate safety briefings based on Mobile VR, and the results we obtained support the effectiveness of this approach.

Self-efficacy results included a surprising detail. Since gaining experience in performing a given behavior is a major factor that increases self-efficacy (Bandura, 1997), one would expect that after successfully donning the life preserver in the real-world, a participant's self-efficacy would increase. Interestingly, although both groups were eventually able to don the life preserver, the further increase in self-efficacy occurred only in the group that used the Mobile VR tool. We speculate that this might be due to the fact that the experiences of donning the life preserver in the Mobile VR world and in the real world were consistent; the real-world experience confirmed and reinforced what users had perceived in the Mobile VR world. In contrast, the experience of passively viewing the safety briefing card pictorials was too different from the act of donning the life preserver, and participants may have been less sure about having correctly executed the donning procedure in the real world.

A potential limitation of the study lies in the nature of the subject sample. The instructional media were evaluated with young participants (age range: 20 to 24, M=21.21, SD=0.64), and people in this age range are likely to be more familiar with using computer applications on touchscreens. While this might be a confound with regard to familiarity with interactive technology vis-à-vis older persons who are generally not as tech savvy, the principles of engagement and interactivity should not be obviated. As such, it would be interesting and informative to repeat the study with people who are less familiar with the technology, and/or belong to other age groups (older adults, children), to determine more fully the generalizability of these results.

Similarly, participants in this study were not well travelled (the mean number of flights in the last two years was only 1.68, SD=2.54) and, therefore, had little exposure to typical pre-flight safety demonstrations on aircraft that can be delivered by flight attendants either live or through a pre-recorded airline video. This inevitably biases the study towards occasional flyers. As such, it would be interesting and informative to repeat the study with people who fly frequently and might have seen flight attendants demonstrate the task multiple times.

A difference between our study and the real-world experience of boarding a commercial aircraft is that in the latter case passengers are exposed to the above mentioned pre-flight demonstration. This situation could be seen as reducing the ecological validity of the evaluation, although the thrust of the study in this paper was not to model a pre-flight safety briefing on aircraft, but to compare two instructional media for efficacy and effectiveness, leaving comparisons of pre-flight safety briefings with the app as a logical next step.

A possible disadvantage of using only Mobile VR as instructional media to provide aviation safety instructions is the following. If the interactive media is part of the aircraft IFE system, a possible technical fault in the IFE system would make the safety media unavailable, and the aircraft would not be allowed to fly. To prevent such situation, the safety briefing card in the seat pockets should remain available to passengers as a backup system. They could also be used by passengers who are not familiar with interactive technology and prefer the traditional solution. For these reasons, the Mobile VR tool is an effective instructional media that can be used by passengers instead of the safety briefing card, but the latter cannot be eliminated, at least in the short run.

After evaluating the Mobile VR tool, as described above, we extended the app with a few additional features. First, we introduced the option of going through the donning procedure without time feedback, which could allow users to familiarize and reflect about the procedure untimed. Second, we added the possibility to share the result obtained in the timed training on public world leader boards. This often leverages social dynamics for users who like competitions. The leader board feature could indeed be an incentive to rehearse the procedure repeatedly, improving the ability to recall it without hesitations. Third, we included an additional 3D world that focuses on the life preserver alone (without the passenger and the cabin environment), giving the user the possibility to examine all its details closely. We made the resulting, extended tool available for download as an app, called 'Life Vest App", for all major mobile platforms (Android, Apple iOS, Windows Mobile). Interested readers can freely try it first-hand on their smartphones or tablets, by downloading it from Google Play (HCI Lab, 2017a), App Store (HCI Lab, 2017b), and Microsoft Store (HCI Lab, 2017c). At the time of writing, the app has been downloaded and installed by 82,000 users (57% Android, 28% Windows Mobile, 15% iOS). Such public accessibility of the app will allow us to conduct remote evaluations, involving large, international, and varied samples of players, to study how they use the Mobile VR tool in more naturalistic ways than laboratory studies.

Finally, we are working at applying the Mobile VR approach described in this paper to other aviation safety topics. In particular, we are currently focusing on two other manual procedures that passengers should learn well to increase their chances of surviving aircraft emergencies: donning the oxygen mask and opening different types of emergency doors on an aircraft.

Acknowledgements

This research was supported in part by a grant from the US Federal Aviation Administration (FAA).

References

Andreatta P. B., Maslowski E., Petty S., et al. Virtual reality triage training provides a viable solution for disaster preparedness. *Academic emergency medicine: journal of the Society for Academic Emergency Medicine* 17(8), pp. 870–876, 2010.

Bandura, A. *Self-efficacy: The exercise of control*. New York, NY, USA: Freeman, 1997.

Bandura, A. Social cognitive theory: An agentic perspective, *Annual Review of Psychology* 52, pp. 1–26, 2001.

Bertram J., Moskaliuk J., Cress U., Virtual Training: Making Reality Work? *Computers in Human Behavior* 43, pp. 284-292, 2015.

Carpenter S.K., Testing Enhances the Transfer of Learning, *Current Directions in Psychological Science* 21, pp. 279-283, 2012.

Chang Y., Liao M., The effect of aviation safety education on passenger cabin safety awareness, *Safety Science* 47, pp. 1337-1345, 2009.

Chang Y., Yang, H. Cabin safety and emergency evacuation: Passenger experience of flight CI-120. *Accident Analysis and Prevention* 43, pp.1049–1055, 2011.

Cha M., S. Han, J. Lee, B. Choi. A virtual reality based fire training simulator integrated with fire dynamics data, *Fire Safety Journal* 50, pp. 12–24, 2012.

Cherry R.G.W., A Study Analyzing the Trends in Accidents and Fatalities in Large Transport Airplanes, *Technical Report DOT/FAA/TC-13/46*, Federal Aviation Administration, Washington, DC, 2013.

Chittaro L., Designing Serious Games for Safety Education: Learn to Brace vs. Traditional Pictorials for Aircraft Passengers, *IEEE Transactions on Visualization and Computer Graphics* 22, pp. 1527-1539, 2016.

Chittaro L., A comparative study of aviation safety briefing media: card, video, and video with interactive controls, *Transportation Research Part C: Emerging Technologies* 85, pp. 415-428, 2017.

Chittaro L., Buttussi F., Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety, *IEEE Transactions on Visualization and Computer Graphics* 21, pp. 529–538, 2015.

Chittaro L., Ranon R., Web3D technologies in learning, education and training: Motivations, issues, opportunities, *Computers and Education* 49, pp. 3-18, 2007.

Cohen, B. H., *Explaining Psychological Statistics* (4th ed.). New York, NY, USA: John Wiley & Sons, pp. 565-566, 2013.

Cohen D., Sevdalis N., Taylor D., et al. Emergency preparedness in the 21st century: training and preparation modules in virtual environments. *Resuscitation* 84, pp. 78–84, 2013.

Corbett C.L., McLean G.A. Passenger safety awareness Reprise 2007: Still ignorant after all these years. Proceedings of the Fifth Triennial International Aircraft Fire and Cabin Safety Research Conference, Atlantic City, NJ, 2007.

Corbett C.L., McLean G.A., Cosper D.K., Effective Presentation Media for Passenger Safety I: Comprehension of Briefing Card Pictorials and Pictograms, *Final Report DOT/FAA/AM-08/20*, Federal Aviation Administration, Washington, DC, 2008.

Corbett C.L., Weed D.B., Ruppel D.J., Larcher K.G., Mc Lean G.A., Inflatable Emergency Equipment I: Evaluation of Individual Inflatable Aviation Life Preserver Donning Tests, *Final Report DOT/FAA/AM-14/14*, Federal Aviation Administration, Washington, DC, 2014.

Cosper D., McLean G.A., Availability of passenger safety information for improved survival in aircraft accidents, *Technical Report DOT/FAA/AM-04/19*, Federal Aviation Administration, Washington, DC, 2004.

Dorsey D., Campbell G., Russell S., Adopting the instructional science paradigm to encompass training in virtual environments, *Theoretical Issues in Ergonomics Science* 10(3), pp. 197–215, 2009.

Edwards M. Stress, behavior, training and safety. *Cabin Crew Safety* 25 (3), Flight Safety Foundation, pp. 1-5, 1990.

Fogg B.J., *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann, San Francisco, 2003.

Grabowski A., Jankowski J. Virtual Reality-based pilot training for underground coal miners, *Safety Science* 72, pp. 310-314, 2015.

Grau R.G., Martínez I.M., Agut S., Salanova M., Safety Attitudes and Their Relationship to Safety Training and Generalised Self-Efficacy, *International Journal of Occupational Safety and Ergonomics* 8(1), pp. 23-35, 2002.

Guo Li H., Chan G., Skitmore M. Using game technologies to improve the safety of construction plant operations, *Accident Analysis and Prevention* 48, pp. 204– 213, 2012.

Harper H., Hedberg J., Wright R., Who benefits from virtuality? *Computers and Education* 34, pp. 163–176, 2000.

HCI Lab, University of Udine. Life Vest App for Android, [Online]. Available: https://play.google.com/store/apps/details?id=it.uniud.hcilab.LifeVestApp [Accessed: 21-October-2017].

HCI Lab, University of Udine. Life Vest App for iOS, [Online]. Available: https://itunes.apple.com/us/app/life-vest-app/id947800480?l=it&ls=1&mt=8 [Accessed: 21-October-2017].

HCI Lab, University of Udine. Life Vest App for Windows Mobile, [Online]. Available: https://www.microsoft.com/en-us/store/p/life-vest-app/9nblggh0930q [Accessed: 21-October-2017].

Katz-Navon T., Naveh E., Stern Z., Safety self-efficacy and safety performance, *International Journal of Health Care Quality Assurance* 20(7), pp. 572-584, 2007.

Li Q., Tay R. Improving drivers' knowledge of road rules using digital games, *Accident Analysis and Prevention* 65, pp. 8-10, 2014.

Luszczynska A., Schwarzer R., Multidimensional health locus of control: Comments on the construct and its measurement, *Journal of Health Psychology*, 10, pp. 633-642, 2005.

Muir H., Thomas L., Passenger education: past and future, *Proc. 4th Triennial Int. Aircraft Fire and Cabin Safety Research Conference*, Lisbon, Portugal, 2004.

National Transportation Safety Board (NTSB). Safety Study: Emergency evacuation of commercial airplanes, *Report NTSB/SS-00/01*, Washington, DC, 2000.

National Transportation Safety Board (NTSB). Accident Report: Loss of thrust in both engines after encountering a flock of birds and subsequent ditching on the Hudson River, *Report NTSB/AAR-10/03*, Washington, DC, 2010.

Seneviratne D., Molesworth B.R.C. Employing humour and celebrities to manipulate passengers' attention to pre-flight safety briefing videos in commercial aviation, *Safety Science* 75, pp. 130-135, 2015.

Schwarzer R., Jerusalem M., Generalized Self-Efficacy scale. In: Weinman, J., Wright, S., Johnston, M. (eds.) *Measures in health psychology: A user's portfolio. Causal and control beliefs*, Windsor, UK: Nfer-Nelson (1995) 35-37

Smith S., Ericson E. Using immersive game-based virtual reality to teach fire-safety skills to children. *Virtual Reality* 13(2), pp. 87–99, 2009.

Stone R.J, Caird-Daley A., Bessell K. SubSafe: A Games-Based Training System for Submarine Safety and Spatial Awareness, *Virtual Reality* 13, pp. 3-12, 2009.

Thomas L.J., Passenger Attention to Safety Information, in *Passenger Behaviour*, R. Bor, ed., Ashgate, UK, 2003.

Weed D.B., Corbett C.L., McLean G.A., Effective Presentation Media for Passenger Safety: Comprehension of Briefing Card Content, Perspective, and Presentation, *Proceedings Seventh Triennial International Aircraft Fire and Cabin Safety Research Conference*, Philadelphia, NJ, 2013.

Zyda M. From visual simulation to virtual reality to games, *IEEE Computer* 38(9), pp. 25-32, 2005.

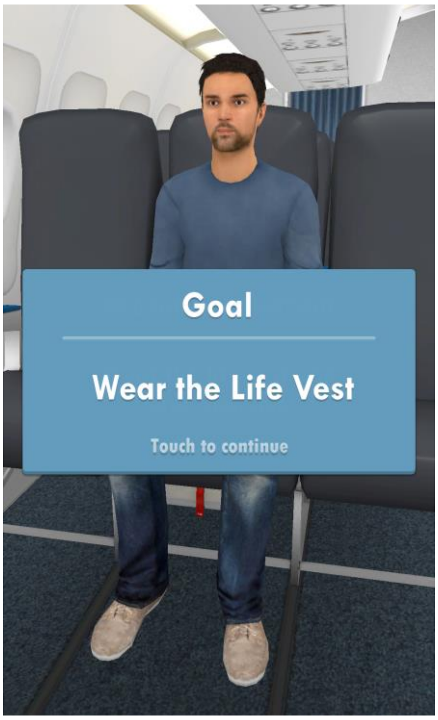

Fig. 1. The virtual passenger in the 3D world.

(a) (b) (c) (d)

Fig. 2. Brief tutorial that instructs the user about how to interact with the 3D world.

Fig. 3. Sequence of steps to don the life preserver.

(a) (b)

Fig. 4. (a) after 6 seconds of user's inactivity, a light bulb icon appears; (b) if the user touches the icon, the app displays a tip about how to proceed.

Fig. 5. The virtual character congratulates the user for the achievement.

Fig. 6. Life preserver compartment under the seat.

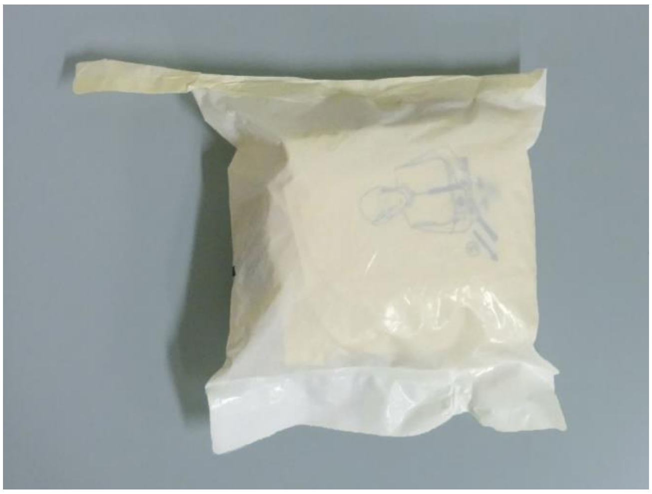

Fig. 7. Pouch containing the life preserver.

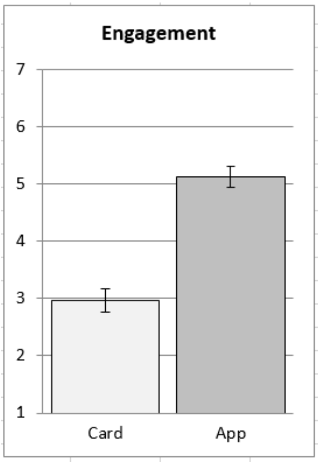

Fig. 8. Engagement with the two instructional media. Capped vertical bars denote ±1 SE.

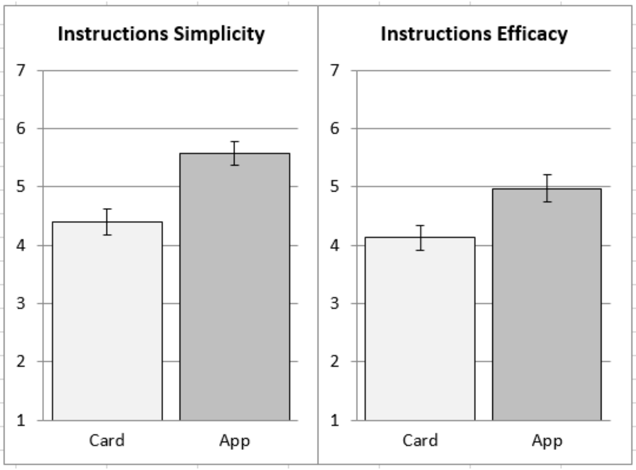

Fig. 9. Perceived simplicity and efficacy of the instructions. Capped vertical bars denote ±1 SE.

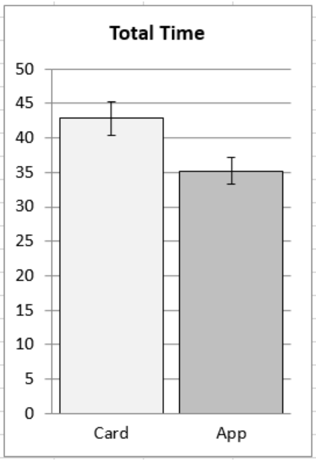

Fig. 10. Total time to perform the donning procedure in the real world. Capped vertical bars denote ±1 SE.

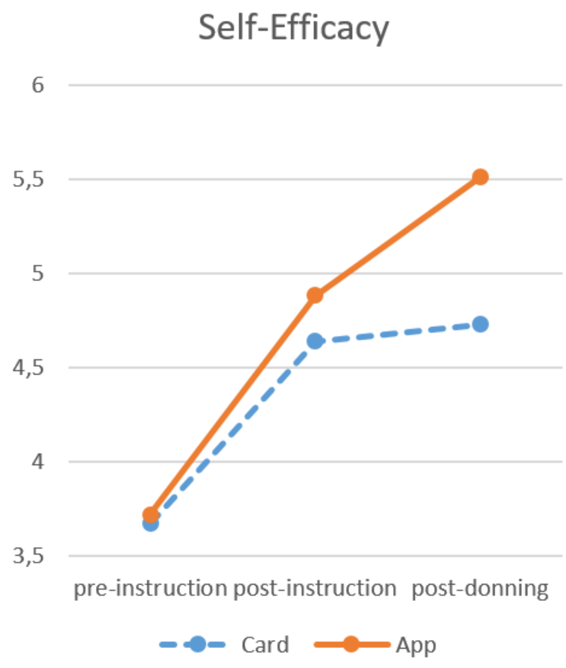

Fig. 11. Self-efficacy measured at three different times: before using the instructional media (pre-instruction), after using the instructional media (post-instruction), and after donning the real life preserver (post-donning).

Table 1.

Number of participants who made an error for each step of the procedure, and description of the errors made. The last line of the table reports the total number of errors made in each group.

Step	Error Made	Card	App
1) Open Compartment	none	0	0
2) Pull out Pouch	none	0	0
3) Open Pouch	trying to open the bottom instead of the top of the pouch	4	0
4) Pull out/Unfold Vest	none	0	0
5) Slip Vest over Head	turning the life preserver more than once to figure out if the two sides are different and/or donning the vest with the buckle on their back	14	6
6) Pass around Strap	trying wrong manouvres with the strap before realizing how to pass it correctly	2	1
7) Buckle Strap	none	0	0
8) Tighten Strap	omission of the step	7	5
	Total Errors	**27**	**12**