Improving Knowledge Retention and Perceived Control through Serious Games: a Study about Assisted Emergency Evacuation

Improving Knowledge Retention and Perceived Control through Serious Games: a Study about Assisted Emergency Evacuation

Luca Chittaro

Abstract— Digital games for education and training, also called serious games (SGs), have shown beneficial effects on learning in several studies. In addition, some studies are suggesting that SGs could improve user's perceived control, which affects the likelihood that the learned content will be applied in the real world. However, most SG studies tend to focus on immediate effects, providing no indication on knowledge and perceived control over time, especially in contrast with nongame approaches. Moreover, SG research on perceived control has focused mainly on self-efficacy, disregarding the complementary construct of locus of control (LOC). This paper advances both lines of research, assessing user's knowledge and LOC over time, with a SG as well as traditional printed materials that teach the same content. Results show that the SG was more effective than printed materials for knowledge retention over time, and a better retention outcome was found also for LOC. An additional contribution of the paper is the proposal of a novel SG that targets the inclusivity goal of safe evacuation for all, extending SG research to a domain not dealt with before, i.e. assisting persons with disabilities in emergencies.

Index Terms— educational games, user interaction, user study, disability, training, safety

1 INTRODUCTION

Digital games for education and training, also called serious games (SGs), have shown beneficial effects on learning in several research studies conducted in different domains [18][16]. In addition, some studies are suggesting that SGs could improve user's perceived control, which affects the likelihood that the learned content will be actually applied in the real world. Indeed, individual differences in perceived control are related to success or failure in a variety of life domains [52].

However, there are important aspects in both the learning and the perceived control lines of research that are still insufficiently explored and need further study. Considering learning, most current studies of SGs focus on immediate learning, providing no indication about retention of the learned content over time. Only a few SG studies [1][11][20][29][46][56] contrasted retention over time of a SG vs. traditional educational materials, with four of them reporting that the SG produced a level of retention superior to the traditional materials. Considering perceived control, SG research has mostly focused on the construct of self-efficacy, which can be defined as the confidence of the individual in his/her ability to perform a behavior [3][4]. However, self-efficacy captures only one of the different aspects of perceived control in Skinner's taxonomy [52]. Another fundamental construct is locus of control (LOC), which can be defined as the degree to which an individual perceives that the outcomes of the situations (s)he experiences are under his/her personal control, or under the control of external factors, such as fate, chance or the actions of other persons [47]. To the best of our knowledge, only one study [10] has focused on the immediate effects of a SG on LOC, suggesting that SGs might make people feel more in personal control of the outcomes of situations they do not feel confident about, which are frequent in health and safety. This motivates the need for further, more thorough explorations of LOC in the SG context, which should go beyond immediate effects on LOC.

This paper aims at advancing the body of knowledge on SGs, through the following contributions. First, it considers a domain that has never been addressed before by SGs, and for which there is a need to improve perceived control as well as promote learning in the general public. The considered knowledge concerns how to help persons with disabilities in evacuating buildings during emergencies. Diffusion of assisted evacuation knowledge in the general public would help in achieving the inclusivity goal of safe evacuation for all [5]. To this purpose, this paper proposes a SG that we designed, implemented and made freely available to the general public in on-line stores. Second, it illustrates a study of the proposed SG that: (i) measured knowledge retention after one week, in addition to immediate learning, (ii) contrasted the SG and traditional pictograms with text, i.e. the media typically used to teach people about emergency evacuations. Third, the study included an investigation of perceived control, focused on the LOC construct.

• *L. Chittaro is with the Human-Computer Interaction Lab, University of Udine, via delle Scienze 206, 33100 Udine, Italy E-mail: luca.chittaro@uniud.it*

TABLE I
CATEGORIZATION OF PREVIOUS STUDIES ABOUT KNOWLEDGE RETENTION OVER TIME

Study	Compared conditions	Topic	Retention Findings
Buttussi et al. (2013)	no comparisons (SG condition only)	Medicine: advanced life support	after 3 months, acquired knowledge was largely retained
Chittaro and Buttussi (2015)	SG vs. traditional printed materials	Emergency preparedness: aircraft cabin safety	after 1 week, knowledge retention with the SG was full and higher than the traditional media
Semeraro et al. (2017)	no comparisons (SG condition only)	Medicine: cardiopulmonary resuscitation	after 3 months, acquired skills were largely retained
Buttussi and Chittaro (2018)	SG played on three different types of display: lower-fidelity headset vs. higher-fidelity headset vs. computer screen	Emergency preparedness: aircraft cabin safety	after 2 weeks, knowledge acquired through the SG was retained equally well regardless of type of display used
Tubelo et al. (2019)	SG vs. traditional printed materials	Medicine: primary health care	after 4 weeks, same level of retention in the SG and traditional media
Din and Gibson (2019)	Computer version of the SG vs. paper version of the SG vs. traditional lecture	Occupational safety: constructions	after (approximately) 2 weeks, higher retention of knowledge acquired through the SG computer version than paper version and traditional lecture
Hu et al. (2021)	Traditional lecture with vs. without the addition of the SG	Medicine: Neonatal resuscitation	after 6 months, retention was higher when the SG was added to the lecture than lecture only
Rahouti et al. (2021)	SG vs. traditional slide-based lecture	Emergency preparedness: fire safety	after 3 weeks, knowledge retention with the SG was higher than the traditional slide-based lecture
Hu, Xiao, and Li (2021)	SG vs. traditional lecture	Medicine: COVID-19	after 5 weeks, knowledge retention was higher with the SG than the traditional lecture
Alrehaili and Al Osman (2022)	SG played on headset vs. SG played on computer screen vs. traditional printed materials	7th grade education: behavior of bees	after 1 week, slightly higher retention of knowledge acquired with the SG on computer screen than traditional media, no significant retention differences between types of display

The paper is organized as follows. In Section 2, we introduce previous SG research on knowledge retention and perceived control. In Section 3, we illustrate in detail the proposed SG. Section 4 and 5 respectively present the study and its results, while Section 6 discusses the results and Section 7 concludes the paper, outlining future work.

2 RELATED WORK AND MOTIVATIONS

The number of SG studies in the literature has reached the hundreds, and a large number of studies has been the object of meta-analyses [6][16][18] that have indicated a positive effect of SGs on knowledge acquisition as well as user's engagement. This section focuses on two aspects that are very rarely investigated in the available studies, i.e. knowledge retention and perceived control over time.

2.1 Effects of Serious Games on Learning and Retention

A limitation of many SG studies is that they measure the learning effects of playing the SG, without including a comparison with traditional, nongame approaches. When such comparison is included, game conditions tend to significantly enhance learning relative to nongame conditions and the effect varies across game design aspects such as game mechanics, visual and narrative features [16]. Unfortunately, the practical implications of this result are limited by the fact that most existing SG studies measure only immediate learning [6][16][18]. They fail to consider retention of the learned content over time, a fundamental factor in deciding which method to adopt in real-world education interventions. Indeed, to empower learners, interventions should help them acquire knowledge and skills that can be retained also after a period of nonuse, minimizing the natural phenomenon known as knowledge and skill decay [2]. Therefore, there is a research need for studies that measure both immediate learning and retention over time, with both game and nongame conditions.

Investigation of knowledge retention over time with SGs is in its infancy, and the number of studies is small (see Table 1). The most addressed domain of retention studies in the SG area is medicine [7][29][30][51][56]. Other topics that have been considered concern aviation [8][11], firefighting [46], occupational safety in constructions [20], and children education [1]. In most studies, the minimum number of participants assigned to a group ranged between 24 and 40 [7][8][11][20][30][46], while it was respectively 50 and 58 in [51] and [29], and only 10 and 13 respectively in [1] and [56]. Seven of the ten studies focus on retention at the temporal scale of weeks. The remaining three studies focus on a larger scale (months), but unfortunately two of them [7][51] do not evaluate the SG against nongame conditions, and the third [30] evaluates the SG only as an adjunct to a traditional lecture, not as a standalone condition. This makes the three studies less informative in understanding what type of educational media can better mitigate knowledge decay.

Six of the ten studies include a comparison of retention with a standalone SG versus traditional, nongame learning conditions: four of them [11][20][29][46] report supe-

rior retention with the SG compared to the other conditions, while one [1] reports only a slighter advantage of the SG, and one [56] does not find a significant difference. This inconsistency in the results can have different explanations that might be related to the topic taught, game design, or study design, e.g. the two studies that found a small [1] or not significant difference [56] were those in which the minimum number of participants assigned to a group was smallest (only 10 and 13, respectively).

It is interesting to note that two [1][8] of the ten studies focus on assessing if playing a SG more immersively with virtual reality headsets could affect outcomes, compared with less immersive displays. Buttussi and Chittaro [8] found that retention with the SG was equally good regardless of the type of display used, and the display significantly affected only player's sense of presence. Similarly, Alrehaili and Al Osman [1] did not find statistically significant differences between headset and computer screen. Although the results of the two studies suggest that the display on which the SG is played might not be a relevant factor for knowledge retention, more studies are clearly needed and this could be one of the different research lines to be developed in SG retention research.

Overall, the majority of the six studies that compared a standalone SG with nongame conditions suggest that SGs might produce a level of retention greater than traditional education materials. SGs might thus be a promising new tool for public educational campaigns, especially if the games are available on web sites and app stores, making it easier for users to play them at any time and place. The public SG we propose and evaluate in this paper is aimed at educating the general public about how to help people with disabilities in emergency evacuations, a domain never addressed before in SG research.

2.2 Effects of Serious Games on Perceived Control

More than fifty years of research demonstrated that perceived control is important to psychological functioning, and individual differences in perceived control are related to success or failure in a variety of life domains [47]. This led researchers to propose different control-related constructs, see [52] for a taxonomy. Of these constructs, locus of control, originally proposed in Rotter's pioneering work [48], and self-efficacy, developed by Bandura [3][4], are very frequently cited and capture two fundamental perspectives on perceived control.

Rotter's Locus of control (LOC) can be generally defined as the degree to which an individual believes that (s)he has control over the outcome of events in life. Given a specific situation, an individual's LOC can have an internal orientation (the individual perceives that she can exert control over the outcome of the situation) or an external orientation (the individual perceives that the outcome of the situation is due to external factors, such as fate, chance, or the actions of other individuals).

Bandura's self-efficacy can be instead defined as the confidence of the individual in his/her ability to perform a behavior. Therefore, while Rotter's construct stresses the importance that individuals believe that outcomes can be influenced by their behaviors, Bandura's construct stresses the importance that individuals believe they are capable of producing the required behaviors [52].

Several studies have shown how both constructs are predictors of actual behavior in different domains: Bandura [4] illustrates how different people with similar skills may perform differently depending on their level of self-efficacy, while Reich and Infurna [47] highlight how locus of control has an impact in several domains such as clinics, economics, health, and business. Since the domain of our SG concerns safety, it is worth noting that both constructs have been shown to be predictors of proper safety behaviors. Indeed, the literature describes positive associations between self-efficacy and safety, see [24][37] for summaries. Research about LOC in risky situations showed instead that internal LOC is a predictor of safer attitudes and behaviors. For example, Wuebker showed that internally-oriented workers have less occupational accidents than externally-oriented ones in industry [59] and hospitals [38]. Similar conclusions have been reached with civil aviation pilots [34][60] and military aviators [35]. Hoyt [28] showed that car passengers with an internal orientation are more likely to wear seat belts, while Montag and Comrey [44] related drivers' internal LOC with safer driving, and a change in drivers' LOC was shown to predict change in safe driving behavior [31].

Given the importance of perceived control, research has focused on identifying antecedents of control that can influence such perception in individuals [47][52]. In recent years, SGs have been considered as a potential antecedent of perceived control, but surprisingly SG studies have concentrated mostly on self-efficacy, inexplicably disregarding LOC. Studies on the possible effect of SGs on self-efficacy [8][12][14][15][19][32][33][43][46][55] agree that a SG can improve player's self-efficacy concerning the behaviors taught. Four of these studies included a comparison with nongame conditions, suggesting that SGs can produce a larger self-efficacy improvement than traditional classroom instruction [33], e-learning [19], printed pictograms [12], and slide-based lectures [46].

To the best of our knowledge, the only SG study that focused on LOC is reported in [10], and concerned a SG that taught the brace position passengers have to assume in aircraft emergency landings. The study considered two groups of 24 users each: one group learned the brace position with the SG, the other with a traditional safety card pictorial as those used by airlines. Learning results indicated a superiority of the SG over the safety card. LOC results showed improvement after playing the SG with an increase in internal orientation and a decrease in external orientation. On the contrary, there was no change in LOC in the safety card group. In other words, participants felt the outcomes of an emergency landing were more under their control after using the SG, not the safety card. Although these results are promising, there is a clear need for more SG studies that might confirm them, also considering other domains. Moreover, it must be noted that the study in [10] assessed only immediate effects on learning and LOC, and not their retention over time. Finally. other facets of LOC are worth investigating. The study in [10] assessed LOC following Rotter's original Internal and

External scales. However, it must be noted that more recent LOC research has adopted finer distinctions. In particular, Levenson [41][42] has substituted External LOC with two separate components. Her approach measures LOC as the extent to which respondents believe outcomes of the situations they encounter are the result of their own actions (Internal scale), the acts of other individuals (Others scale), or simply luck (Chance scale). It would thus be interesting to study the effects of SGs also using Levenson's more detailed assessment.

3 The Proposed Serious Game

The need for creating novel ways of educating the general public about how to help persons with disabilities in evacuations emerged in a collaborative project involving our university, the regional medical emergency services, the firefighter corps, and several associations of persons with disabilities (see the Acknowledgment section for the full list). Initial meetings with first responders and the associations revealed that available education and training initiatives are based on non-interactive materials, and typically aimed at professional responders instead of the general public. We considered the development of a SG for the potential of making learning of assisted evacuation knowledge more attractive, engaging, comprehensible, and easy to remember. These are essential features for an intervention aimed at the general public but could benefit first responders as well. We decided that the SG should be capable of running on mobile devices to maximize its availability to users, since the number of smartphones in use significantly surpasses that of PCs, and at the end of 2020 was three times larger [49].

To the best of our knowledge, the proposed SG is the first to deal with teaching how to help disabled persons in evacuations. The SG has been publicly released with the name *Help! – The serious game*, and is freely available on both Android [26] and iOS [27]. As of June 2023, its number of installations is 35000 (68% Android, 32% iOS). The language of the characters' voices and the text in the game can be selected between English and Italian.

3.1 Knowledge acquisition process

The knowledge taught by the game was obtained and validated through a 2-stage process. In the first stage, we examined general documents about first response in the disability context, produced by different organizations such as the National Department of Firefighters in Italy [21] and the Center for Development and Disability in the US [9], and we had meetings with one of the authors of the Italian guidelines. The analysis of such documents allowed us to identify a set of basic rules that any citizen should be aware of, because they can make a difference in guiding his/her behavior during an emergency that involves persons with disabilities. In the second stage, we conducted meetings with the associations of persons with disabilities to discuss with them the knowledge we planned to include in the SG. This stage was especially useful to more thoroughly understand social aspects of the interaction with the person in the emergency, for example how to properly address the person, and actions/words to avoid because they would make the person uncomfortable. Such aspects were thus also included in the SG.

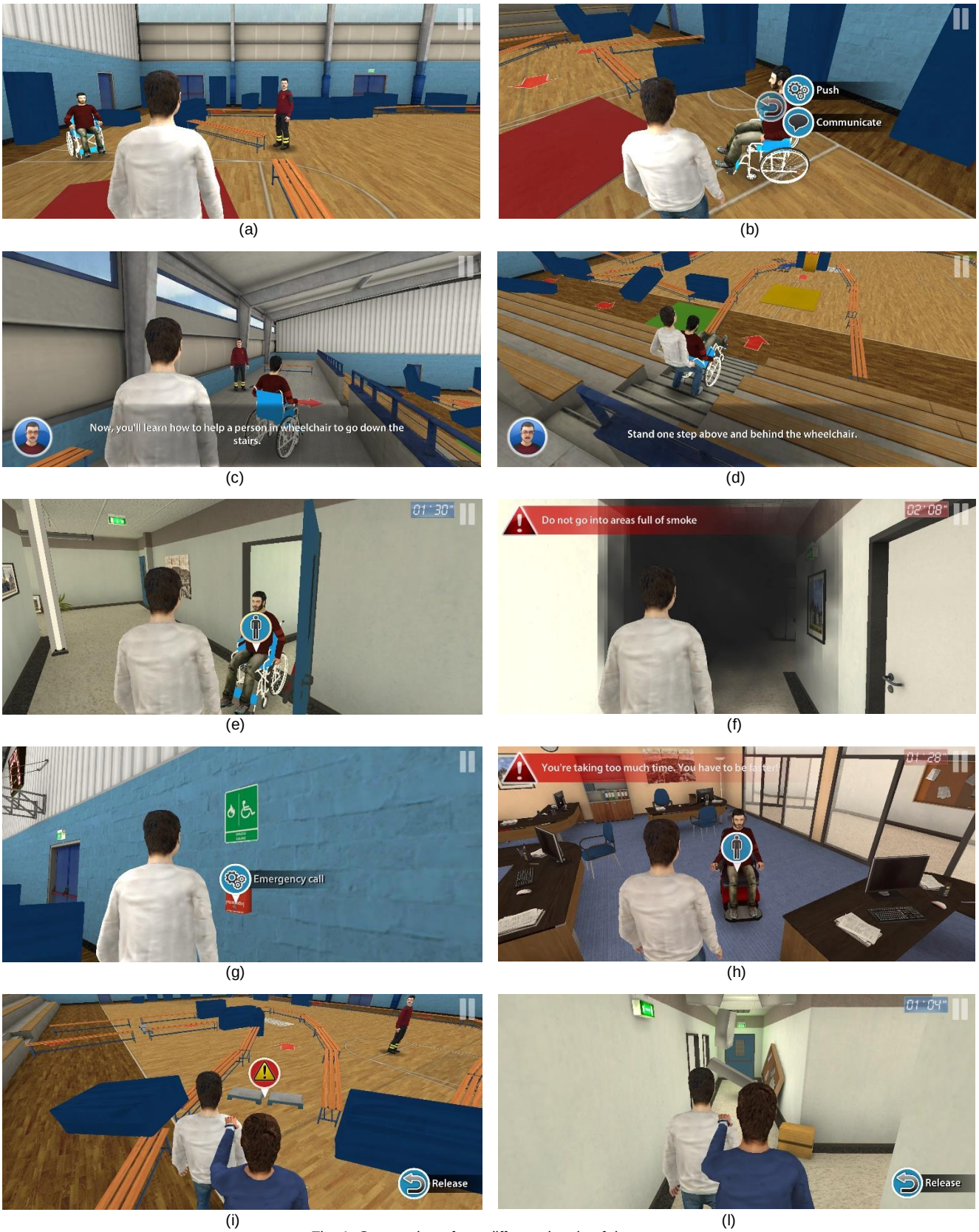
3.2 Game design

We initially organized the SG in 8 different levels (5 dealing with motor disability, 3 with vision disability). In the game, players see their avatar in third-person view as in Figure 1 (the player's avatar is the character wearing a white sweatshirt). To control avatar movement, we adopted the familiar type of interaction found in many mobile games: players can move their right thumb on the right part of the touchscreen to change avatar orientation and associated viewpoint, and the left thumb on the left part of the touchscreen to make the avatar move forward/backward or strafe left/right. When actions concerning objects or characters in the environment are available, an icon appears on the object or character (e.g., Fig. 1b, 1e, 1g, 1h, 1i). Touching the icon allows the player to perform the action. When there is more than one action available, the touched icon expands into more icons as shown in Fig. 1b, and the player can touch which action to perform. In all levels, if the player performs an action that is inappropriate at a given moment, the SG plays an error sound, flashes a red warning sign and explains the error with a short text message, e.g., Fig. 1f.

The first level of the SG is set in a large gym (Fig. 1a, 1b, 1c, 1d) where the player encounters an animated pedagogical agent (APA) representing a firefighter, and another character (a person in a wheelchair). The APA speaks to the player to teach him/her the different types of actions (s)he should perform to effectively help the person in an evacuation. The voice of the APA is also subtitled (e.g., Fig. 1c, 1d). After each explanation of an action, the APA invites the player to perform it. For example, after explaining how the wheelchair should be pushed down the stairs, the player should actively push it (e.g., Fig. 1d), and the motion of the chair down the stairs reflects its real-world slow speed to prepare the user about what to expect in an actual evacuation.

After the APA has presented all the knowledge concerning motor disability scenarios, and the player has executed all the actions correctly, the next four levels are set in different types of emergencies (two earthquakes of different strength, two fires of different size) and involve different types of motor disability, two requiring a manual wheelchair (e.g., Fig. 1e), two a powered wheelchair (e.g., Fig. 1h), and different types of environmental threats such as smoke (e.g., Fig. 1f), heavy obstacles (e.g. Fig. 1l), risk of falling objects (e.g, Fig. 1l), broken glass, risk of electric shock (e.g. Fig. 1e), to be handled properly by the player.

In the sixth level, the player is brought back to the training gym with the APA, but this time the third character has a visual disability (e.g., Fig. 1i), and the APA teaches how to handle evacuation of blind persons. Then, the next two levels are set in different types of emergencies, earthquake (e.g., Fig. 1l) and fire, and involve different types of threats, but this time the player should apply the knowledge concerning how to help blind persons.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (l)

Fig. 1. Screenshots from different levels of the game.

While the levels depicting emergencies are played, a watch in the top right of the screen (e.g., Fig. 1e, 1f, 1h, 1l) shows time taken, which is an important factor in evacuations. If the evacuation (or an action of the evacuation) is taking a long time, the game warns the user about it (e.g., Fig. 1h). Moreover, if the player ignores threats in the emergency or does not handle them appropriately, the SG plays an error sound, flashes a red warning sign and ex-

TABLE 2
ITEMS OF THE LOCUS OF CONTROL QUESTIONNAIRE

SUBSCALE	**If I find myself in an emergency evacuation with a person with a disability, succeeding in helping the person…**
INTERNAL	… depends mostly on my ability
CHANCE	… is controlled by accidental happenings
OTHERS	…does not depend on me, but on firefighters and medical emergency services
CHANCE	… is mostly a matter of luck
OTHERS	…would be unlikely without the help of first responders
INTERNAL	… is determined by my own actions
CHANCE	… is a matter of fate
OTHERS	…would not end well if first responders were late
INTERNAL	… depends on how hard I try
OTHERS	…depends mostly on the person with a disability
CHANCE	… depends on whether I'm lucky enough to be in the right place at the right time

plains the issue with a short text message (e.g., Fig. 1f).

4 USER STUDY

The between-groups study we conducted aimed at assessing and comparing the effects of two different ways of teaching the considered knowledge: a group of participants used traditional printed materials (PM), i.e. pictograms with text, while another group used the SG.

4.1 Participants

The study involved 60 participants (28 F, 32 M), with various occupations. They were volunteers recruited through personal contact, and did not receive any compensation. All participants were regular smartphone users, and were able to use mobile applications. We asked them to rate how familiar they were with helping persons with disabilities on a 7-point scale (1=not at all, 7=very). Participants were assigned to the two groups in such a way that: (i) each group had 30 participants (SG: 13F, 17M; PM: 15F, 15M); (ii) the two groups were similar in terms of age (SG: M=27.97, SD=12.30; PM: M=28.87, SD=8.70) as well as familiarity with helping disabled persons, which was very low (SG: M=2.17, SD=1.53; PM: M=2.53, SD=1.46). Each of these two variables was submitted to a one-way ANOVA that confirmed lack of statistically significant differences between the groups. Frequencies of gender in the two groups were instead analyzed with a 2x2 chi-square test showing that there was no significant association between gender and group.

4.2 Materials

The SG group used the game described in Section 3 on an Android smartphone in Full HD resolution. Since participants were Italian, language in the game settings was set to Italian. To prevent the introduction of confounding factors, the pictograms used in the PM group were created by taking screenshots of the game, and the text included in the pictograms was all and only the same text used in the SG group to describe what to do in assisted evacuation. In this way, we avoided the confounding variables that are found in studies that employ printed materials with a graphical appearance and/or textual content different from the game. Concepts and actions were also presented in the same sequence of the game. This resulted in a set of 58 pictograms, each printed to appear in the same size of the screen used by the SG group. The printed document had a total of 15 pages, A4 size.

4.3 Measures

4.3.1 Knowledge

To measure participants' prior knowledge, how much it changed after exposure to the material, and how much was retained after 1 week, we used 13 questions about the different elements that both educational materials taught. To avoid suggesting possible answers (e.g., as in a multiple-choice questionnaire), we asked participants to answer the questions orally and we recorded the answers. Some questions asked to define a single concept (e.g., what is an evacuation chair?) or take a single choice about a situation (e.g., when should you push the wheelchair to help the person in the evacuation?), and were scored 1 if the answer was correct, 0 otherwise. Other questions asked to list multiple items (e.g., what threats do you have to stay away from during the evacuation?) or the different actions needed to achieve a goal (e.g., how do you help a person in a wheelchair to go down the stairs?), and the score was the number of correct items (or actions) that the participant was able to remember. The total score of the knowledge test was the sum of the scores of the individual questions, and could range from 0 to 31.

4.3.2 Locus of Control

To measure participants' LOC concerning helping persons with disabilities in emergencies, we adapted items from Levenson's IPC scale [42] to our context, by introducing in them explicit mentions to evacuation aspects. We devoted four items to each of the three subscales, which respectively deal with the respondent's perception of how much (s)he is responsible of the outcome of the emergency evacuation (INTERNAL subscale), how much the outcome depends on chance factors (CHANCE subscale), and how much it depends on other people (OTHERS subscale). We asked participants to indicate their level of agreement with each of the 12 items on a 6-point scale (1=strongly disagree, 6=strongly agree). Table 2 shows the association between each questionnaire item and the corresponding subscale. For each subscale, answers to the four items were averaged to form a reliable scale. Cronbach's alpha at pre-test, post-test and retention-test was: .57, .67, .74 (INTERNAL); .71, .72, .81 (CHANCE); .58, .60, .69 (OTHERS).

4.3.1 EGameFlow

EGameFlow [22] is an instrument that measures different aspects of the user experience of educational games, and is organized into 8 subscales. One of the subscales concerns social interaction among players through the

game and thus does not apply to our case, because our SG is single-player. We thus administered the other seven subscales (a total of 34 items) to participants who used the SG. The subscales respectively concern: i) concentration, i.e. how much the participant can remain concentrated on the game, ii) goal clarity, i.e. how much the goals of the game are clear, iii) feedback, i.e. quality of perceived feedback provided by the game, iv) challenge, i.e. how appropriate is the level of challenge and its increase as the participant progresses in the game, v) autonomy, i.e. sense of control on the game perceived by the participant, vi) immersion, i.e. how much the participant feels immersed in the game, vii) knowledge improvement, i.e. how much the participant feels that the game helps improving his/her knowledge. EGameFlow asks participants to rate each item on a 7-point scale (1=strong disagreement, 7=strong agreement).

4.4 Hypotheses

Considering the results about knowledge of the three studies in the literature that compared retention with SGs vs. traditional printed materials, involving 24 or more participants per group [11][29][46], we expected to obtain, also in our different domain, a knowledge increase in both groups but a superior retention over time in the SG group compared to the PM group. Regarding LOC, the only study of LOC effects of a SG [10] found that the SG was superior to pictogram-based printed materials in improving Internal and External LOC, but evaluated the effect only immediately after exposure and not over time. Considering the above mentioned promising evidence about knowledge retention, we expected better results on Internal LOC also over time. Since we adopted Levenson's approach that does not consider Internal and External LOC as mutually exclusive dimensions (in which the increase of the first leads to a decrease of the second), and substitutes External LOC with the Chance and Others components, our study is exploratory in nature with respect to these two components. Consideration of EGameFlow data is also exploratory, to highlight possible correlations between specific aspects of the user experience and possible improvements in knowledge or LOC.

4.5 Procedure

Participants were informed that we were testing different materials that illustrate how to help persons with disabilities in emergency evacuations. The experimenter obtained consent for participation, and for audio recording of answers to the knowledge test. Participants were informed that they could refrain from continuing their participation in the study at any time and without the need to provide a reason. Then, the demographic questionnaire, the LOC questionnaire, and the knowledge test, were administered. After the pre-test measurements, participants in the PM group received the printed materials and were asked to examine the content of all the pages completely and without time limitations. Participants in the SG group were given the smartphone with the SG and were asked to play all the levels completely and without time limitations. When the experimental condition was completed, we administered again the LOC questionnaire and the knowledge test to all participants. Participants in the SG group filled also the EGameFlow (before the LOC questionnaire, so that the game experience was as close in time as possible to the EGameFlow measurement). Participants were thanked and asked for their availability to answer some additional questions one week later, without specifying what those further questions were going to concern. After one week, participants were administered again the LOC questionnaire and the knowledge test.

5 Results

For measures assessed over time (knowledge and LOC), we analyzed results using a 2x3 mixed design ANOVA, in which group served as the between-subjects variable and time of measurement served as the within-subjects variable. Effect sizes are reported as partial eta squared (η_p^2). Following [17], when a statistically significant interaction between group and time was found, we analyzed the effects of time separately for each group with repeated measures ANOVA, and the effects of group separately at each time of measurement with between subjects ANOVA.

5.1 Knowledge

Since Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2)=18.81$, $p<.001$), degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon=.781$). Statistically significant results of the 2x3 ANOVA revealed a main effect of time, $F(1.561, 90.549)=450.38$, $p<.001$, $\eta_p^2=.89$, and a group by time interaction effect, $F(1.561, 90.549)=19.27$, $p<.01$, $\eta_p^2=.11$. Mean values are shown in Figure 2. For PM, they were 4.80 (SD=2.12) at pre-test, 17.17 (SD=4.01) at post-test, and 15.23 (SD=4.33) at retention test. Mean values for SG were 3.60 (SD=1.75) at pre-test, 18.87 (SD=4.20) at posttest, and 17.57 (SD=3.91) at retention-test. Investigation of the interaction revealed a statistically significant effect of time on knowledge in both groups (PM: $p<.001$, SG: $p<.001$). Pairwise comparisons with Bonferroni test were statistically significant between pre-test and post-test (PM: $p<.001$; SG: $p<.001$), pre-test and retention-test (PM: $p<.001$; SG: $p<.001$), post-test and retention-test (PM: $p=.001$; SG: $p<.05$). The difference in knowledge between groups was statistically significant at pre-test ($p<.05$) and retention-test ($p<.05$).

5.2 Locus of Control

Statistically significant results of the 2x3 ANOVA for Internal LOC revealed a main effect of time, $F(2,116)=27.92$, $p<.001$, $\eta_p^2=.33$, and a group by time interaction effect, $F(2,116)=3.23$, $p<.05$, $\eta_p^2=.05$. Mean values are shown in Figure 3. For PM, they were 4.55 (SD=.61) at pre-test, 4.99 (SD=.74) at post-test, and 4.81 (SD=.78) at retention test. Mean values for SG were 4.44 (SD=.71) at pre-test, 5.33 (SD=.63) at post-test, and 4.99 (SD=.88) at retention-test. Investigation of the interaction revealed a statistically significant effect of time on knowledge in both groups (PM: $p<.01$, SG: $p<.001$), and pairwise comparisons with Bonferroni test were statistically significant

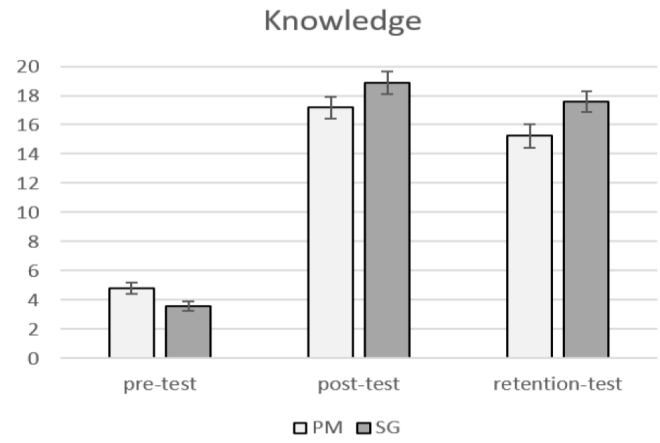


Fig. 2. Means of knowledge. Capped vertical bars indicate ± SE.

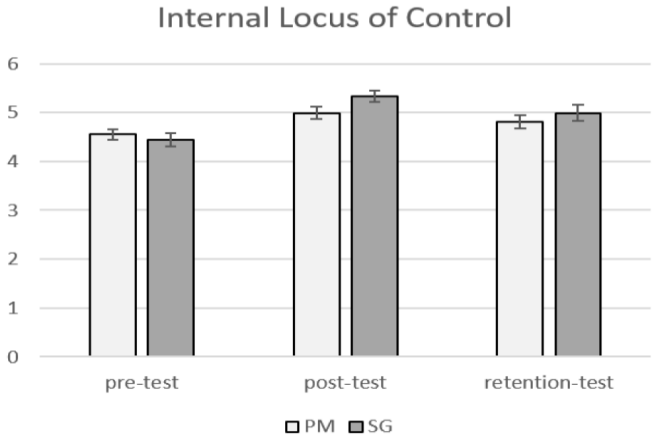


Fig. 3. Means of Internal LOC. Capped vertical bars indicate ± SE.

between pre-test and post-test in both groups (PM: $p<.05$; SG: $p<.001$), while they were significant only in the SG group between pre-test and retention-test (SG: $p=.001$) and between post-test and retention-test (SG: $p<.05$). The 2x3 ANOVA did not instead reveal any statistically significant effects for the Chance and Others subscales.

5.3 EGameFlow

The EGameFlow results (Figure 4) were the following: i) Concentration, M=5.91, SD=.94, ii) Goal Clarity, M=6.23, SD=.97, iii) Feedback, M=6.30, SD=.79, iv) Challenge, M=5.02, SD=1.38, v) Autonomy, M=5.49, SD=1.13, vi) Immersion, M=4.11, SD=1.43, vii) Knowledge Improvement, M=6.16, SD=.83. We analyzed the possible correlations between the perceived characteristics of the user experience and the changes in knowledge and in Internal LOC that occurred in players between pre-test and post-test, and between post-test and retention-test. Statistically significant results indicated that perceived Challenge, Immersion, and Knowledge Improvement positively correlated with the difference in knowledge observed between pre-test and post-test. The correlation was medium for Challenge ($\rho=.39$, $p<.05$), and large for Immersion ($\rho=0.58$, $p=.001$) and Knowledge Improvement ($\rho=0.59$, $p=.001$). Since players' knowledge increased between pre-test and post-test, the three constructs are related to a larger increase in knowledge between pre-test and post-test.

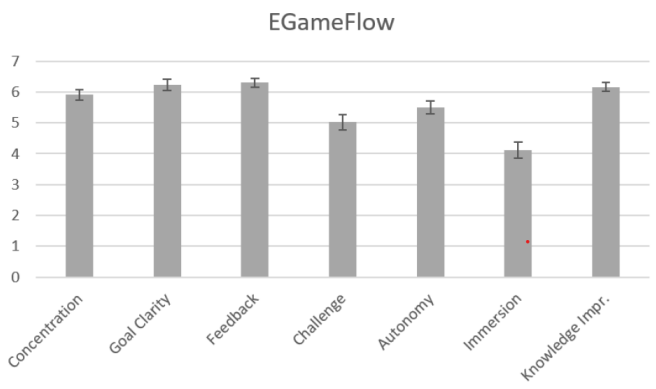


Fig. 4. Means of EGameFlow. Capped vertical bars indicate ± SE.

Challenge, Immersion, and Goal Clarity positively correlated with the change in Internal LOC that occurred in players between post-test and retention-test (Challenge: $\rho=0.46$, $p=.01$; Immersion: $\rho=0.37$, $p<.05$: Goal Clarity: $\rho=0.38$, $p<.05$). Since player's Internal LOC decreased between those two times, the three constructs are related to a smaller decrease (that is, a larger retention) of Internal LOC between post-test and retention-test.

6 Discussion

Participants in both groups had statistically significant greater knowledge at post-test rather than pre-test and, although they both suffered a natural knowledge decay after one week, it was small and their knowledge remained significantly greater at retention-time than pre-test. However, participants who had used the SG retained a greater level of knowledge than those who used printed materials, confirming our hypothesis on the positive role of the SG on knowledge retention. The fact that participants in the SG group started at a statistically significant disadvantage in knowledge at pre-test time, with slightly less knowledge than the PM group, further highlights the greater effectiveness of the SG over printed materials. Knowledge results confirm our hypothesis, reinforcing the conclusions of [11][29][46], and extending them to a new domain. Overall, knowledge results suggest that advantages of SGs in learning become more evident over time, with a mitigation of the natural phenomenon of knowledge decay [2], which is important for real-world interventions. Unlike other studies, the comparison with the nongame condition in our study was designed in a highly controlled way, as described in section 4.2. As a result, the factor that played a role in determining the better performance of the SG is the difference between the two types of media (PM and SG) that was allowed by our design, i.e. the interactive nature of the relationship between user and material. Participants in the PM group were indeed exposed to a set of printed pictograms that covered all the same topics of the SG by employing exactly the same graphics and text of the SG (obtained by taking screenshots of the game) and following the same sequence of presentation of the SG. Participants in the SG group actively controlled one of the characters in that same graphical environment (with the addition of sound)

and received the same text messages. It is worth noting that isolating contributing factors is more difficult in the other studies that compared retention with a SG and a nongame condition, because they did not have this level of control, and thus their experimental conditions included a larger number of differences, which can confound results. More specifically, the study in [11] compared a small set of pictograms with a SG that did not match the pictogram sequence, and contained a more extensive narrative and content (relayed through voice and text feedback); in [1], participants in one group read text in a booklet, while the other group did something very different (performing flights as a bee in a virtual environment that kept text short and concise); in [56], one group read a (2600-word) booklet, while the other group had to consider clinical cases with fictional characters in the SG and diagnose them; in [29], one group listened to a university lecture, while the other group screened and attended patients in the SG; Rahouti et al. [46] also compared a (slide-based) lecture with a SG in which participants navigated a visual replica of their real working place and followed a storyline.

Regarding perceived control, participants in both our groups had a statistically significant greater Internal LOC at post-test rather than pre-test. However, after the 1-week decay that occurred, only the SG group, and not the PM group, was able to retain a statistically significant advantage in perceived control at retention-test with respect to pre-test. It might also be worth mentioning that the difference in Internal LOC between the two groups at post-test time approached significance ($p=.059$). The positive effect of the SG on Internal LOC is consistent with the only other study of a SG on LOC [10]. However, that study considered LOC only immediately after playing, while we extended the result by showing that improvement in LOC persists after 1 week only in the SG group, a more interesting finding for real-world interventions. There were instead no statistically significant effects concerning the Chance and Others subscales. Unlike Internal LOC, this result cannot be compared to [10] because that study relied on External LOC as a single construct in Rotter's tradition, while we used Levenson's two-components measure of external factors (see Sections 2.2 and 4.4).

The obtained LOC results suggest that SGs could help improve Internal LOC outcomes of an educational intervention. It would be interesting to identify which features of the SG gave it a superior capability with respect to printed materials. As mentioned before, the difference between the PM and the SG condition in our study consisted in interactivity. However, some features of interactivity could be more important than others. The literature on the relationships between SGs and self-efficacy has argued that enactive learning, supported by the interactivity of SGs, is an important mechanism that can explain how playing a SG can result in self-efficacy increase [39]. According to Social Cognitive Theory [3][4], in enactive learning the participant learns from the consequences of his/her actions, and when they end up in success, the resulting mastery experience promotes self-efficacy. Although our study concerned locus of control, not self-efficacy, it must be noted that the two constructs capture two different and fundamental aspects of the higher-level construct of perceived control [47][52]. It is thus reasonable to hypothesize that enactive learning and mastery experiences can be in the same way an important mechanism behind the positive effect of the SG on Internal LOC. More specifically, actively controlling the character can emphasize players' ability in choosing and performing, in the virtual world of the game, the behaviors that determine the successful outcome in the considered situations, while printed materials do not allow to do that. Moreover, some authors who studied self-efficacy have stressed that self-identification with the SG character could be an especially important mediator in the relationship between interactivity and self-efficacy [40][45]. Future studies on LOC with SGs could include measurement of self-identification, e.g. by using the Player-Avatar Identification scale [40]. If a significant role emerged, it could become worthwhile to introduce avatar personalization features in the SG with the aim of improving positive effects on LOC.

Regarding the correlations between the EGameFlow subscales and the changes in knowledge and Internal LOC of players, the first two reported correlations concerning knowledge might suggest a relevance of Challenge and Immersion of the game experience, while the third correlation shows that participant's subjective perception of being helped by the game in learning correlated with actual knowledge increase, suggesting an awareness in players of the positive effect of playing the game on knowledge.

The correlations concerning Internal LOC confirm the special role of Challenge and Immersion, and might suggest a greater relevance of those two features of the game experience, together with Goal Clarity, for the desired outcome on player's LOC. Overall, Challenge and Immersion were the two aspects related to both knowledge and Internal LOC outcomes. In a systematic review of SG studies [36], challenge came out as the most commonly cited aspect that could promote engagement and learning. It is worth noting that the Immersion construct of the EGameFlow is instead different from the immersion and presence constructs [50][53] often used in the literature about 3D environments. Immersion in EgameFlow refers to how much players become involved in the game, and lose track of time and space [23]. The correlation obtained aligns thus to the discourse on learner's engagement in SGs and its possible positive outcomes in education [36].

Regarding the specific educational content, which concerned assisted evacuation, a limitation of the study is that it considered knowledge about motor and vision disabilities, but not hearing and cognitive disabilities. We are currently working on hearing disabilities, and the public release of the SG has been extended with three levels that begin to address it, bringing the total number of levels to 11. However, teaching how to communicate in sign language with a deaf person in a SG is technically more complex to implement because sign language is not universal, but country-dependent. Therefore, unlike the levels described in this paper, which are universally valid

and could be easily subtitled and/or dubbed in any language, a different level with different animations should be built for each sign language in the world. The three hearing disability levels we have added to the game currently contain animations in one sign language, but there are 300 sign languages in the world [57].

6 Conclusions

This paper contributed to research about learning and retention with SGs by showing that a SG was more effective than traditional printed materials to promote knowledge retention over time, reinforcing conclusions of previous studies and extending them to a domain that had not been addressed before, i.e. helping persons with disabilities in emergencies. To this purpose, the paper proposed and described a novel SG that we made available for download to the general public [26][27]. A third contribution of the paper is that we extended the investigation to possible changes in user's locus of control as a result of playing the SG. This topic had been considered by only one study, and that investigation was limited to immediate effects without considering retention. Our results confirmed superiority of the SG over printed materials also on LOC outcomes over time. In addition to the on-going extension to hearing disability (sketched in the previous section), our future work will consider cognitive disabilities. In particular, emergencies are a notable example of uncertain experiences in which cognitive disability can lead to unpredictable and often counterproductive behavior that rescuers have difficulty in understanding and managing. An interactive simulation could be an ideal educational tool to familiarize with such circumstances and how to handle them.

A promising line of research would be to study the effect of repetitive playing: while in our case participants played the game just once, one could conjecture that repetitive playing over time could both lead to greater gains in knowledge and LOC and to better consolidation of the gains over time. Such repetitive exposure research could also explore another potential benefit of SGs: while people could be less motivated to periodically re-examine printed materials, a SG that is able to provide a good level of engagement might make them more willing to rehearse the procedures periodically in the game environment.

To extend the scope of the proposed SG as a tool to teach assisted evacuation, we are also considering the inclusion of new scenarios with further environmental and operational barriers. In particular, assisted evacuation of aircraft would be an especially relevant and challenging scenario. In previous work [13], we created a SG to teach how to deal with aircraft emergencies involving passengers without disabilities, and we could build on that experience to introduce different aircraft types as new environments of the assisted evacuation SG. However, the development of this new content should take into account expert guidance that is currently being drafted by aviation regulatory authorities [22]. More generally, another possible development could be the inclusion of the presence of a crowd in enclosed environments: this would be useful to introduce issues such as exit visibility, social influence and bounded rationality, e.g. [25] [58].

The study conducted in this paper could be extended by collecting naturalistic data online from a larger sample of players who have installed the SG on their devices. Following the approach described in [13], the analysis of wrong and right actions of a very large sample of players could allow to better assess learning needs of the general public as well as to study competence gain over time and identify which features of the game experience are more or less effective. In addition, recording the movement of players in the environments reproduced by the game would allow us to feed that data to deep learning systems that are used with real-world movement data, see e.g. [54], to predict and simulate players' behavior.

Finally, the possibility of personalizing characters' appearance mentioned in the discussion could be explored up to the level of creating a digital twin of the player. This would support two different research directions. First, we could use the player's digital twin as his/her controlled character that helps the person with disabilities, and assess if this can lead to improvements in perceived control outcomes with respect to using a generic character as the current game does. Second, the player's digital twin could be instead used to represent the person with disabilities that the player has to help, and we could assess if this can lead to improvements in player's empathy towards persons with disabilities.

Acknowledgments

This work was partially supported by the Friuli Venezia Giulia region under the project "Servizi avanzati per il soccorso sanitario al disabile basati su tecnologie ICT innovative" ("Advanced emergency medical services for the disabled based on innovative ICT technologies"). In defining the content of the proposed game, we received precious information and feedback from the following Italian organizations: Association of the Blind and Visually Impaired (UICI), Association of the Deaf and Mute-Deaf (ENS), regional autism association (PROGETTO AUTISMO FVG), Dystrophy Association (UILDM), Regional Council of Persons with Disabilities (CONSULTA FVG), Association of Quadriplegic and Paraplegic People, Stroke Association (ALICe), National Association of Visually Impaired People's Families (ANFaMiV), Association of the Families of Persons with Disabilities (GE.CO.).
Nicola Zangrando (HCI Lab, University of Udine) and Samantha Modolo contributed to the development and testing of the game. We also thank Stefano Zanut (Italian Firefighters Corps) for sharing his experience on defining national guidelines for assisted evacuation.

References

[1] E.A. Alrehaili, and H. Al Osman, "A virtual reality role-playing serious game for experiential learning", *Interact. Learn. Environ.*, vol. 30, pp. 922-935, 2022.

[2] W. Arthur, E. A. Day, W. Bennett, and A. M. Portrey, eds., *Individual and team skill decay: The science and implications for practice*. Routledge/Taylor & Francis Group, 2013.

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

[4] A. Bandura, "Social cognitive theory: An agentic perspective", *Annu. Rev. Psychol.*, vol. 52, pp. 1-26, 2001.

[5] K. Boyce, "Safe evacuation for all - Fact or Fantasy? Past experiences, current understanding and future challenges", *Fire Saf. J.*, vol. 91, pp. 28-40, 2017.

[6] E.A. Boyle, T. Hainey, T.M. Connolly, G. Gray, J. Earp, M. Ott, T. Lim, M. Ninaus, C. Ribeiro, and J. Pereira, "An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games", *Comput. Educ.*, vol. 94, pp. 178-192, 2016.

[7] F. Buttussi, T. Pellis, A. Cabas Vidani, D. Pausler, E. Carchietti, and L. Chittaro, "Evaluation of a 3D serious game for advanced life support retraining", *Int. J. Med. Inform.*, vol. 82, pp. 798–809, 2013.

[8] F. Buttussi, and L. Chittaro, "Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario", *IEEE Trans. Vis. Comput. Graph.*, vol. 24, pp. 1063-1076, 2018.

[9] Center for Development and Disability, *Tips for First Responders, 5th edition*, https://unmhealth.org/services/development-disabilities/programs/other-disability-resources/first-responder-tips.html (last accessed 16 June 2023)

[10] L. Chittaro, "Designing Serious Games for Safety Education: 'Learn to Brace' versus Traditional Pictorials for Aircraft Passengers", *IEEE Trans. Vis. Comput. Graph.*, vol. 22, no. 5, pp. 1527-1539, 2016.

[11] L. Chittaro, and F. Buttussi, "Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety", *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 4, pp. 529-538, 2015.

[12] L. Chittaro and F. Buttussi, "Exploring the use of arcade game elements for attitude change: Two studies in the aviation safety domain", *Int. J. Hum. Comput.*, vol. 127, pp. 112-123, 2019.

[13] L. Chittaro, and F. Buttussi, "Learning Safety through Public Serious Games: A Study of 'Prepare for Impact' on a Very Large, International Sample of Players", *IEEE Trans. Vis. Comput. Graph.*, vol. 28, pp. 1573-1584, 2022.

[14] L. Chittaro, and R. Sioni, "Serious games for emergency preparedness: Evaluation of an interactive vs. a non-interactive simulation of a terror attack", *Comput. Hum. Behav.*, vol. 50, pp. 508-519, 2015.

[15] E.J.M. Chee, L. Prabhakaran, L.P. Neo, G.A.C. Carpio, A.J.Q. Tan, C.C.S. Lee, and S.Y. Liaw, "Play and Learn with Patients-Designing and Evaluating a Serious Game to Enhance Nurses' Inhaler Teaching Techniques: A Randomized Controlled Trial", *Games Health J.*, vol. 8, no. 3, pp. 187-194, 2019.

[16] D. B. Clark, E. E. Tanner-Smith, and S. S. Killingsworth, "Digital games, design, and learning a systematic review and meta-analysis", *Rev. Educ. Res.*, vol. 86, pp. 79–122, 2016.

[17] B.H. Cohen, *Explaining Psychological Statistics*, 4th ed., New York, NY, USA: John Wiley & Sons, 2013.

[18] T. M. Connolly, E. A. Boyle, E. MacArthur, T. Hainey, and J. M. Boyle, "A systematic literature review of empirical evidence on computer games and serious games", *Computers & Education*, vol. 59, pp. 661–686, 2021.

[19] M.E.W. Dankbaar, O. Richters, C.J. Kalkmanet, et al., "Comparative effectiveness of a serious game and an e-module to support patient safety knowledge and awareness". *BMC Med. Educ.*, vol. 17, 2017.

[20] Z.U. Din, and G.E. Gibson, "Serious games for learning prevention through design concepts: An experimental study", *Saf. Sci.*, vol. 115, pp. 176-187, 2019.

[21] Dipartimento dei Vigili del Fuoco, *Il soccorso alle persone disabili*, http://www.vigilfuoco.it/allegati/biblioteca/legge_disabili.pdf (last accessed 16 June 2023)

[22] Federal Aviation Administration, "Safety Considerations in the Air Transportation of Persons With Disabilities", Draft, *Advisory Circular AC 120-32A*, Washington, DC, 2022.

[23] F. L. Fu, R. C. Su, and S. C. Yu, "EGameFlow: A scale to measure learners' enjoyment of e-learning games", *Comput. Educ.*, vol. 52, no. 1, pp. 101-112, 2009.

[24] R.G. Grau, I.M. Martínez, S. Agut, and M. Salanova, "Safety Attitudes and Their Relationship to Safety Training and Generalised Self-Efficacy", *Int. J. Occup. Saf. Ergon.*, vol. 8, no. 1, pp. 23-35, 2002.

[25] M. Haghani, and M. Sarvi, "'Rationality' in Collective Escape Behaviour: Identifying Reference Points of Measurement at Micro and Macro Levels", *J. Adv. Transp.*, ID 2380348, 2019.

[26] HCI Lab, University of Udine. *Help! – The serious game for Android*, https://play.google.com/store/apps/details?id=it.uniud.hcilab.helpseriousgame (last accessed 16 June 2023)

[27] HCI Lab, University of Udine. *Help! – The serious game for iOS*, https://apps.apple.com/us/app/help-the-serious-game/id1323966262 (last accessed 16 June 2023)

[28] M.F. Hoyt, "Internal-external control and beliefs about automobile travel", *J. Res. Pers.*, vol. 7, pp. 288-293, 1973.

[29] H. Hu, Y. Xiao, and H. Li, "The Effectiveness of a Serious Game Versus Online Lectures for Improving Medical Students' Coronavirus Disease 2019 Knowledge", *Games Health J.*, vol. 10, pp. 139-144, 2021.

[30] L Hu, L. Zhang, R . Yin, Z. Li, J . Shen, H . Tan, J . Wu, and W. Zhou, "NEOGAMES: A Serious Computer Game That Improves Long-Term Knowledge Retention of Neonatal Resuscitation in Undergraduate Medical Students", *Front. Pediatr.*, vol. 9, 645776, 2021.

[31] J. L. Huang, and J. K. Ford, "Driving locus of control and driving behaviors: Inducing change through driver training", *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 15, no. 3, pp. 358-368, 2012.

[32] L. Huang, and Y. Yeh, "Mediated enactive experience versus perceived mastery experience: An enhancing mechanism versus a mediator of character attachment and internal political efficacy in serious games", *Comput. Hum. Behav.*, vol. 55, pp. 1085-1096, 2016.

[33] C.M. Hung, I. Huang, and G.J. Hwang, "Effects of digital game-based learning on students' self-efficacy, motivation, anxiety, and achievements in learning mathematics", *Comput. Educ. J.*, vol. 1, pp. 151–166, 2014.

[34] D.R. Hunter, "Development of an aviation safety locus of control scale", *Aviat. Space Environ. Med.*, vol. 73, pp. 1184-1188, 2002.

[35] D. R. Hunter, and J. E. Stewart, "Safety locus of control and accident involvement among army aviators", *Int. J. Aviat. Psychol.*, vol. 22, no. 2, pp. 144-163, 2012.

[36] A.I. Jabbar, and P. Felicia, "Gameplay Engagement and Learning in Game-Based Learning: A Systematic Review", *Rev. Educ. Res.*, vol. 85, no. 4, pp. 740-779. 2015.

[37] T. Katz-Navon, E. Naveh, and Z. Stern, "Safety self-efficacy and safety performance", *Int. J. Health Care Qual. Assur.*, vol. 20, no. 7, pp. 572-584, 2007.

[38] J.W. Jones, and L.J. Wuebker, "Safety locus of control and employees' accidents", *J Bus Psychol.*, vol. 7, pp. 449-457, 1993.

[39] Y. H. Lee, "Does Digital Game Interactivity Always Promote Self-efficacy?", *Cyberpsychol. Behav. Soc. Netw.*, vol. 18, no. 11, pp. 669-673, 2015.

[40] D.D. Li, A.K. Liau, and A. Khoo, "Player–Avatar Identification in video gaming: Concept and measurement", *Comput. Hum. Behav.*, vol. 29, no. 1, pp. 257-263, 2013.

[41] H. Levenson, "Multidimensional locus of control in psychiatric patients", *J Consult Clin Psychol.*, vol. 41, pp. 397–404, 1973.

[42] H. Levenson, "Differentiating among internality, powerful others, and chance", in *Research with the locus of control construct: Vol. 1. Assessment methods,* H. M. Lefcourt, Ed., New York, NY, USA: Academic Press, pp. 15–63, 1981.

[43] A. Meluso, M. Zheng, H. A. Spires, and J. Lester, "Enhancing 5th graders' science content knowledge and self-efficacy through gamebased learning", *Comput. Educ.*, vol. 59, no. 2, pp. 497-504, 2012.

[44] I. Montag, and A.L. Comrey, "Internality and externality as correlates of involvement in fatal driving accidents", *J. Appl. Psychol.*, vol. 72, pp. 339-343, 1987.

[45] W. Peng, "Design and evaluation of a computer game to promote a healthy diet for young adults", *Health Commun.*, vol. 24, no. 2, pp. 115-127, 2009.

[46] A. Rahouti, R. Lovreglio, S. Datoussaïd, and T. Descamps, "Prototyping and Validating a Non-immersive Virtual Reality Serious Game for Healthcare Fire Safety Training", *Fire Technol.*, vol. 57, pp. 3041–3078, 2021.

[47] J.W. Reich, and F.J. Infurna, Eds., *Perceived Control*, Oxford, UK: Oxford University Press, 2017.

[48] J. B. Rotter, "Generalized expectances for internal versus external locus of control", *Psychol. Monogr.*, vol. 80, pp. 1–28, 1966.

[49] Samsung, https://insights.samsung.com/2021/08/19/your-phone-is-now-more-powerful-than-your-pc-3/ (last accessed 16 June 2023)

[50] M. Sanchez-Vives, and M. Slater, "From presence to consciousness through virtual reality". *Nature reviews: Neuroscience*, vol. 6, pp. 332-339, 2005.

[51] F. Semeraro, A. Frisoli, C. Loconsole, N. Mastronicola, F. Stroppa, G. Ristagno, A. Scapigliati, L. Marchetti, and E. Cerchiari, "Kids (learn how to) save lives in the school with the serious game Relive", *Resuscitation*, vol. 116, pp. 27-32, 2017.

[52] E. A. Skinner, "A guide to constructs of control", *J. Pers. Soc. Psychol.*, vol. 71, no. 3, pp. 549-570, 1996.

[53] M. Slater, and S. Wilbur, "A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments", *Presence*, vol. 6, pp. 603–616, 1997.

[54] X. Song, R. Shibasaki, N. J. Yuan, X. Xie, T. Li, and R. Adachi, "DeepMob: learning deep knowledge of human emergency behavior and mobility from big and heterogeneous data", *ACM Trans. Inf. Syst.*, vol. 35, no. 4, pp. 1-19, 2017.

[55] H. Sung, and G. Hwang, "A collaborative game-based learning approach to improving students' learning performance in science courses", *Comput. Educ.*, vol. 63, pp. 43-51, 2013.

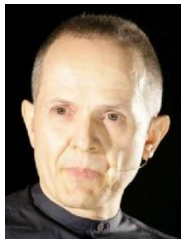
[56] R.A. Tubelo, F.F. Portella, M.A. Gelain, M.M.C. de Oliveira, A.E.F. de Oliveira, A. Dahmer, and M.E.B. Pinto, "Serious game is an effective learning method for primary health care education of medical students: A randomized controlled trial", *Int. J. Med. Inform.*, vol. 130, 103944, 2019.

[57] United Nations, (2022) https://www.un.org/en/observances/sign-languages-day (last accessed 9 September 2022)

[58] L. Wang, and Y. Jiang, "Escape dynamics based on bounded rationality", *Phys. A: Stat. Mech.*, vol. 531, 121777, 2019.

[59] L.J. Wuebker, "Safety locus of control as a predictor of industrial accidents and injuries", *J. Bus. Psychol.*, vol. 1, pp. 19-30, 1986.

[60] X. You, M. Ji, and H. Han, "The effects of risk perception and flight experience on airline pilots' locus of control with regard to safety operation behaviors", *Accid. Anal. Prev.*, vol. 57, pp. 131-139, 2013.



Luca Chittaro is full professor of Human-Computer Interaction (HCI) in the Department of Mathematics, Computer Science, and Physics of the University of Udine, Italy, where he heads the HCI Lab (http://hcilab.uniud.it). He has authored or co-authored over 200 international academic publications, and he has been an ACM Distinguished Speaker. His major research interests are in virtual reality, serious games, persuasive technology, mobile HCI, and their applications in health and safety. He has received research grants from organizations such as the US Federal Aviation Administration (FAA) and the European Union (EU), and companies such as the Benetton Group and the Intesa Sanpaolo banking group.