NOTICE: This is a paper that was accepted for publication in the **IEEE Transactions on Learning Technologies**. It is not the final published version.

The DOI of the definitive version available in IEEE Explore is

https://doi.org/10.1109/TLT.2019.2902401

Humor and Fear Appeals in Animated Pedagogical Agents: An Evaluation in Aviation Safety Education

Fabio Buttussi and Luca Chittaro

Abstract—Humor and fear appeals are widely employed in traditional communication for educational purposes, but their exploitation in animated pedagogical agents has been scarcely explored. We studied the use of humor and fear appeals by a three-dimensional animated pedagogical agent that taught the same procedural knowledge in four conditions: i) humor appeal, ii) fear appeal, iii) humor and fear appeal together, or iv) no emotion appeals. The agent delivered and illustrated the knowledge in an application for aviation safety education. The results of our study show that using the educational application had overall positive learning effects, regardless of the appeals used by the agent. Resort to humor and fear appeals by the agent affected instead participants' mood, liking of the application, and perception of the agent. The paper includes a discussion of the advantages and disadvantages of using humor and fear appeals in animated pedagogical agents.

Index Terms—Animated agent, pedagogical agent, embodied agent, virtual agent, emotion appeal, humor, fear, education, user evaluation, aviation safety

1 Introduction

ANIMATED pedagogical agents (APAs) are lifelike autonomous characters that cohabit learning environments with students to create rich, face-to-face learning interactions [1]. Emotions are increasingly considered as an important factor of APAs. In particular, APAs can express emotions (e.g., in [2], [3], [4]), or consider students' emotions to adapt the educational content in affective tutoring systems (e.g., in [5], [6], [7]).

In traditional education, researchers have been studying the effects of appealing to emotions for more than sixty years. In particular, the use of humor by human educators was found to attract and maintain students' interest and attention [8], relieve tension [9], increase perception of educator's likeability [10], and improve students' motivation [11]. Fear appeals were used especially in health and safety education, and under certain circumstances they were found to be effective in improving attitudes towards and engagement in healthy and safe behaviors [12], [13], [14]. Humor and fear appeals are instead scarcely explored in APAs.

This paper studies the use of humor and fear appeals by a three-dimensional APA that taught the same procedural knowledge in four conditions: i) humor appeal, ii) fear appeal, iii) humor and fear appeal together, or iv) no emotion. The APA is part of an application for aviation safety education. In this education domain, as well as other health and safety education domains (e.g., disease prevention, occupational safety,…), people often tend to underestimate the risks and not to pay the required attention to traditional lessons and educational materials (e.g., [15], [16], [17]).

The paper aims at advancing knowledge in two main directions. On one hand, we want to assess if humor and fear appeals in APAs could affect learning and knowledge transfer. On the other hand, we want to explore the possible effects of the considered appeals on perception of different aspects of the APA, on liking of the application, and on participants' mood, which is a factor that could affect performance of the taught procedures [18]. Notably, this is the first study that considers the use of humor and fear appeals by APAs as two independent variables, analyzing the four possible conditions, including the one in which humor and fear appeals are used together, and assessing the possible interactions between humor and fear appeals.

The paper is organized as follows. Section 2 presents background information about humor and fear appeals in education, and related work about emotions in APAs. Section 3 illustrates the APA we developed. Section 4 describes the study. Section 5 and 6 report and discuss results. Finally, Section 7 concludes the paper, outlining future work.

• *Fabio Buttussi and Luca Chittaro are with the Human-Computer Interaction Lab, Department of Mathematics, Computer Science, and Physics, University of Udine, Italy. E-mail: {fabio.buttussi, luca.chittaro}@uniud.it.*

2 Background and Related Work

2.1 Humor in Education

The use of humor in classroom education has been the object of several studies. The review by Banas et al. [19] defines different types of humor, outlines the main theories and findings from several studies, and provides educators with practical advice about using humor in the classroom. Humor has been defined as the communication of incongruities that are amusing in some manner [20], and the use of both verbal and nonverbal communication to elicit laughter and joy [21]. Torok et al. [22] distinguished between *positive* types of humor (e.g., jokes) and *negative* ones (e.g., aggressive humor), and found that positive types of humor were generally recommended by students, while only sarcasm was recommended among negative ones. Wanzer et al. [23] distinguished between humor that is appropriate in education (*appropriate humor*) and humor that is not (*inappropriate humor*). Banas et al. [19] identified positive types of humor that are always appropriate, negative ones that are always inappropriate, and positive and negative ones whose appropriateness depends on the context. Humor related to class material (*related humor*) was generally considered as the most appropriate one in [24]. Indeed, related humor does not distract students from the instructional message and can make information more memorable [19].

Among the general theories of humor, the Instructional Humor Processing Theory (IHPT) [25] focuses on humor in education. It proposes that students need to perceive and then resolve the incongruity in the humorous message. If the incongruity is not resolved, humor may lead to confusion or distraction. If the incongruity is resolved but humor creates a negative affect, or if it creates a positive affect but it does not enhance the ability to process the educational content, then humor will have negative or no impact on learning. Finally, if humor creates a positive affect and enhances the ability to process the educational content, then it will lead to learning and retention. A study [25] showed that appropriate humor positively correlates with learning, while inappropriate humor does not. Other studies (e.g., [26]) found that students were confused by humor, and others (e.g., [11]) found no effect of humor on learning. Ziv [27] observed that if humor is not related to the taught content, it is not expected to improve learning and may also be remembered better than the content. Moreover, he observed that many studies did not conduct manipulation checks to assess if humor was perceived as funny. To support learning, Ziv suggested this sequence: i) introduce the concept, ii) illustrate it with a joke, and iii) paraphrase the concept after students stop laughing. Different studies showed that the use of humor can increase perception of educator's likeability [10] and credibility [11], but not perceived expertise [10], [11]. However, too much humor can also diminish educator's credibility and the use of negative humor can negatively affect evaluations received by educators [22]. Humor can relieve tension especially on anxiety-provoking topics [9], and frequent short bursts of humor were suggested to attract and maintain students' interest and attention [8].

Although Banas et al. [19] suggested as future research direction the integration of humor into education through online classes and other technologies for learning, very little research has focused on this topic. Houser et al. [11] considered video lessons delivered by means of CD-ROMs, and found that humor had no significant effect on student learning, but it significantly improved students' motivation and perceptions of instructor credibility. Humorous aviation safety education videos used as pre-flight briefings were found to positively affect mood, but greater entertainment was also associated to poorer retention of the safety messages [28]. In the same domain, a serious game based on arcade game elements, including humorous situations with cartoon characters, was found to engage users and increase knowledge about correct and wrong behaviors in aircraft emergencies [29]. A recent study [30] compared traditional lessons about basic concepts of information technology with or without the use of humor and cartoons. Results showed that humor and cartoons improved students' attitudes toward the lesson, decreased anxiety, and positively affected knowledge retention. Unfortunately, there are only a few studies about the effects of humor used by APAs, as we will describe in Section 2.3.

2.2 Fear Appeals in Health and Safety Education

Fear appeals are persuasive messages designed to scare people about risks and negative consequences of specific actions with the aim of changing people's attitudes and behaviors [13], [31], [32]. Fear appeals have been widely used in health and safety education. Examples include educating people about sun protection to prevent skin cancer [14], hygiene procedures to limit the spread of avian influenza [33], proper driving behaviors to avoid car accidents [34], risks of tobacco and alcohol abuse [35], [36], self-protection procedures in terror attacks or aircraft emergencies [37], [38], [39].

To understand when fear appeals are successful or not, psychologists studied how people respond to information and recommendations about risks. In particular, two leading theoretical models, the Extended Parallel Process Model [32] and Protection Motivation Theory [40], point out that people's perception of the risk and the actions suggested to avert it are key factors for the effectiveness of fear appeals. In particular, fear appeals must highlight the *severity* of the risk and the *vulnerability* of the message recipient to the risk. If people perceive the risk as not severe or perceive themselves as not vulnerable to it, then the models predict they will not feel threatened, and will not be motivated to consider how to cope with the risk. Moreover, fear appeals must highlight that there are effective actions to avert the risk (*recommendation efficacy*), and that people are capable of performing such actions (*recommendation simplicity*). If people perceive the actions as ineffective or perceive themselves as not capable of performing them, then the models predict that people will try to reduce negative emotions induced by threat through processes such as risk denial and defensive reactions, which are detrimental to learning proper behavior. Numerous studies investigated the effects of fear appeals presented through print or video advertising (e.g., [14], [35], [36]). Recently,

researchers started to explore the effects of fear appeals in digital media such as Web-based education [33] and serious games [37], [38], [39].

2.3 APAs and Emotions

APAs are the combination of two ideas from previously distinct research areas: animated interface agents for human-computer interaction, and knowledge-based learning environments [1]. The area of animated interface agents devoted considerable effort to embodied conversational agents (ECAs) that aim at having the same properties as humans in face-to-face conversation such as recognizing verbal and non-verbal input, generating verbal and non-verbal output, and giving signals that indicate the state of the conversation [41]. APAs can be created both as ECAs (*conversational APAs*) and animated interface agents that do not engage users in conversations (*non-conversational APAs*).

Emotions in animated agents have been mainly studied in the context of ECAs. Since the early proposals, APAs included support for non-verbal communication that also allowed them to convey some emotions. For example, summarizing previous work, Johnson et al. [1] highlighted the importance of eye-gaze to regulate turn-taking, head nods or shakes to provide positive or negative feedback, and facial expressions such as smiles to indicate agreement or pleasant surprise in response to students finishing their tasks. This section focuses on studies that investigated emotions in the context of APAs. For example, Mao and Li [7] found that APA expressiveness influenced students' satisfaction. Chen et al. [42] found that an empathic APA increased students' willingness to continue reading and complete exercises. Baylor and Kim [4] studied the effects of facial expressions and gestures on both attitudinal instruction (changing students' beliefs regarding intellectual property) and procedural instruction (learning how to perform tasks within a software program). Interestingly, they found that for attitudinal instruction, participants learned more when the APA face expressed emotions but deictic gesture was absent, while for procedural instruction, students learned more when the APA gestures were present. Beale and Creed [43] reviewed studies that focused on affective aspects of APAs. Unfortunately, they found inconsistent results for the educational domain. In [3], APAs that displayed emotions (emotional APAs) were rated more positively than unemotional ones. In [44], APAs with higher levels of emotional intelligence were no better than those with lower levels of emotional intelligence at creating a social relationship with subjects. In [2], an emotional APA was perceived to be less helpful than an unemotional one. In [45], an emotional and an unemotional APA were not different in terms of perceived entertainment, difficulty, and distraction, but the emotional APA was perceived to be more sympathetic and enhanced concentration and motivation.

Very few APAs used elements of humor. For example, the APA in [46] used some quirky asides that were well perceived by users, and the APA in [47] fidgeted when it was dragged across the screen. In [48], an APA with enthusiastic feedback and sense of humor was compared with the same APA without these features, and the results showed a modest increase in users' self-efficacy. The APA in [49] welcomed users with an opening joke, then it taught a history lesson in three parts of 10-15 minutes each. After each part, it produced a humorous performance from a database of humorous riddles and funny pictures. The APA also empathically reacted to users' emotions, which they reported by means of emoticons. The use of humor and empathy was found to improve learning and motivation. Unfortunately, no study considered the use of humor and fear appeals as two independent variables to assess the possible effects and interactions on learning, perception of the APA, and students' mood.

3 The Proposed APA

In this section, we describe a non-conversational APA that can teach procedures using different kinds of emotion appeals. The APA organizes a lesson about a procedure in the following four phases:

1. Introduction. The APA welcomes users and verbally introduces the procedure.
2. Demonstration. The APA demonstrates and verbally describes the sequence of steps the procedure is made of.
3. Practice. The APA asks users to interact with it and/or the virtual environment to perform first-hand each step of the procedure. During this phase, the APA monitors users' interaction, and provides users with feedback. The feedback about each user's attempt to complete a step is provided immediately after the attempt. More precisely, if users complete a step correctly at the first attempt, the APA provides them with a positive feedback about correct step completion. If users fail at the first attempt, the APA tells them that they did not perform the specific step correctly. If users fail at subsequent attempts, the APA demonstrates and verbally explains how to perform the step. If users complete a step correctly after one or more failed attempts, the APA acknowledges step completion.
4. Final feedback. The APA gives a verbal feedback about the entire practice phase in terms of errors and/or time to complete it. More precisely, the APA informs users: i) if they made errors, or ii) if they made no errors, but spent too much time to perform the procedure, or iii) if they completed the practice phase correctly and quickly. After the first practice phase, the APA invites users to try a second time. If users accept, the APA restarts the practice phase. Otherwise, the lesson concludes.

The behavior of the APA at the abstract level above is defined by a set of general if-then rules, that can be applied to any procedure contained in a database. Procedures are defined by specifying the steps they are made of, and the associated APA demonstrative animations, spoken sentences, body language, and facial expressions. In this way, we can make changes limited to the emotion-dependent parts of the database to have the APA teach the same procedures using different kinds of emotion appeals.

Fig. 1. Screenshots of the APA demonstrating how to place A) the feet, B) the head, and C) the hands.

For the study, we developed four variants of a lesson that teaches the brace position, i.e. the position that passengers should assume to protect themselves during an emergency landing. In all four variants, the procedure is made of three steps, demonstrated by the APA: how to place the feet (Fig. 1A), then the head (Fig. 1B), and finally the hands (Fig. 1C). Therefore, during the practice phase, the APA asks users to act on its feet, then its head, and finally its hands, placing them in the correct positions. The four variants differ in the use of emotion appeals, which involves some spoken sentences, body language, and facial expressions used by the APA.

The spoken sentences were recorded in the native language of the study participants (Italian). For reader's convenience, this paper provides their English translation. Since humor is often dependent on cultural context, we will provide the additional contextual details where necessary.

The first variant does not use any emotion appeal. The APA goes through the different phases without displaying any emotion (neither positive nor negative) in its words, voice, body language, and facial expressions (see for example Fig. 2A). The content of each phase is the following:

1. Introduction. The APA welcomes users, says that it will explain an important procedure, and points out that an emergency landing is rare, but requires attention. Then, it says that users should prepare themselves by learning the brace position.
2. Demonstration. The APA demonstrates and describes the three steps of the brace procedure. After showing all the steps, the APA says that the brace position has to be kept during all the landing.
3. Practice. As positive feedback, the APA simply says "Right", "Exact", or "Correct", in a neutral voice. If users make a first error, the APA says, based on the step of the procedure, that the feet, head, or hands are not in the correct position. If users make additional wrong attempts on the same step, the APA demonstrates again the step. As acknowledgements for completing a step, it says "Ok" or "Done", in a neutral voice.
4. Final feedback. Based on how users performed the Practice phase, the final feedback is given through one of the following sentences: "You have made one error in the trial", "You have made some errors in the trial", "You should assume the brace position faster", "You have assumed the correct position quickly, but it is useful to practice", all in a neutral voice.

The second variant uses a fear appeal. Following the theories described in Section 2.2, the APA adds remarks to highlight that users are susceptible to the risk (vulnerability) and that the consequences are severe (severity), but also that the taught procedure is effective to avert the risk (recommendation efficacy) and that users can perform it by following a few simple steps (recommendation simplicity). The APA can use its body language to underline the remarks, and its facial expressions to show concern about the severity of the consequences (see for example Fig. 2B and 2C). The tone of the voice is serious when the APA corrects users' errors, and sad when it mentions death and severe injuries as possible consequences of not complying with the procedure. The content of each phase is the following:

1. Introduction. The APA welcomes users, says that it will explain an important procedure, and points out that an emergency landing is rare, but requires attention. Then, it explains in a sad voice that impact with the ground can be so hard that it can cause severe injuries or even death (Fig. 2B). Finally, it introduces the effective action to avert the risk, saying in a serious voice that to avoid those consequences users must know the brace position.
2. Demonstration. The APA shows and describes the three steps of the procedure as in the first variant, and adds a remark: after placing its head against the seat in front of it, the APA tells users that if they do not perform that step, they will risk a head trauma. This is told while the APA has its head against the seat to show the simple effective action that averts the risk (Fig. 2C). Finally, the APA makes a remark in a serious voice ("If you don't want to suffer fatal hits") before telling users to keep the position throughout the landing.
3. Practice. When the APA gives the positive feedback ("Right", "Exact", or "Correct") to users who correctly complete a step at the first attempt, it also remarks that they would have suffered a serious consequence (specifying which one among broken legs, head trauma, or broken arms) if they would not have placed the involved body part as they have just done. This highlights that a simple action

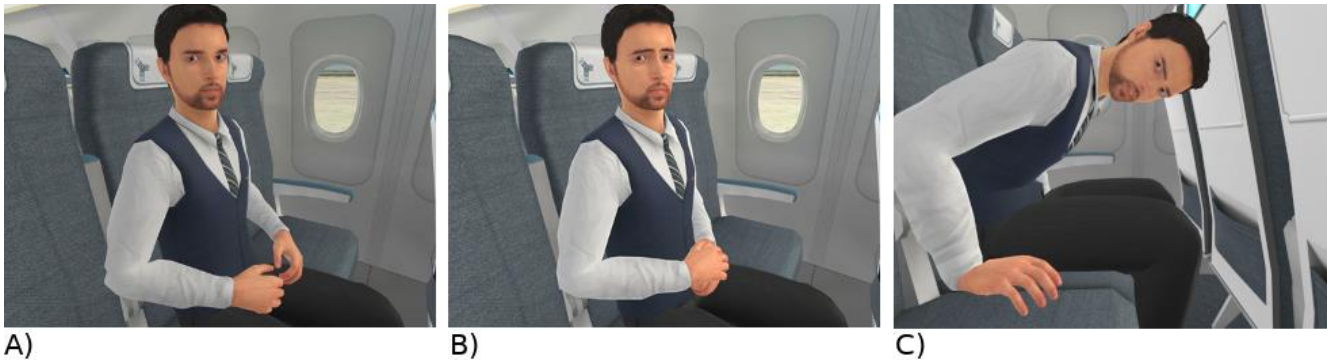

Fig. 2. Screenshots of the APA A) displaying no specific emotion, B) being concerned about the severity of the consequences during introduction, and C) being concerned during the remark about head position.

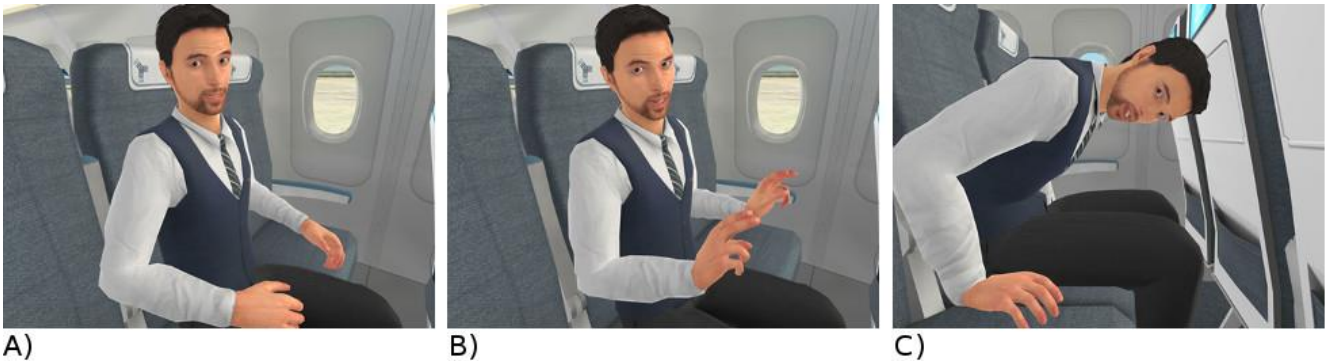

Fig. 3. Screenshot of the APA A) smiling and raising its shoulders, B) drawing quotes in the air with its fingers, and C) pretending to be surprised during the humorous comment about its shoes.

that users have just performed is effective to avert the risk. If users fail to place a body part correctly at the first attempt, the APA says that if they place that body part in that way, they are going to suffer the specific trauma for that body part. In this way, we make sure that users are equally exposed to the fear appeal, regardless of their success or failure at the first attempt. The demonstrations after the subsequent wrong attempts and the eventual acknowledgment of step completion are the same used in the first variant.

4. Final feedback. The APA tells the same sentences used in the first variant, but it adds the following remarks. When the APA says that users have made errors in the practice phase, it adds with a sad voice that if they make errors in the brace position, they may die. After saying that users should assume the brace position faster, the APA remarks that if they are too slow, the impact will happen when they are not in the correct position and it will cause them serious injuries. Finally, when the APA tells users that they have assumed the correct position quickly, it suggests that to avoid severe injuries, it is useful to practice.

The third variant uses a humor appeal. The APA adds humorous comments, puns, and jokes to the procedure presentation and the feedback. Since IHPT (as seen in Section 2.1) posits that recognizing and resolving incongruity is fundamental to positively affect learning [25], the APA uses tone of voice, body language, and facial expressions to underline the use of humor (see for example Fig. 3) and make incongruity easier to recognize. Moreover, since [25] showed that inappropriate humor has negative or no impact on learning, humor used by the APA does not fall in the categories of inappropriate humor identified by [19] and [23]. In particular, since [19] and [24] found that humor related to the class material is generally the most appropriate one, humor used by the APA focuses on the steps of the procedure or the feedback that it gives to users. It is important to note that in this variant the APA uses humor in the same circumstances in which it uses fear in the second variant, so the number of times the APA aims at evoking emotions in the two variants is the same. The content of each phase is the following:

1. Introduction. The APA welcomes users, says that it will explain an important procedure, and points out that an emergency landing is rare, then it adds a humorous quote which is well-known in the cultural context of the participants' country, making the humor easily recognizable. The popular quote can be translated in English as "But, you know, good luck is blind, but bad luck sees very well". The APA underlines that it is a humorous comment with tone of voice, and by smiling and raising its shoulders (see for example Fig. 3A). Then, when the APA invites users to pay attention and tells them to prepare themselves by learning the brace position, it adds a pun saying that the "brace position" is technically called "let's hug ourselves tightly". This pun is particularly related to the content and is meant to enhance the ability to process the message, which has a key role in IHPT [25], because the brace position can be remembered as the act of embracing one's own legs. The APA emphasizes the pun by drawing

quotes with its fingers as shown in Fig. 3B, and uttering the words with a funny accent.

2. Demonstration. The APA shows and describes the three steps of the procedure as in the first variant, and adds the following joke. After placing its head against the seat in front of it (second step), the APA tells users not to be afraid to ruffle their hair, and it adds that they will receive a complimentary comb for each emergency landing. The joke is told when the APA has its head against the seat, and the humorous reference to the hair may help users remember that their head should be against the seat. Finally, when the APA says to keep the position during all the landing, it describes the position as "very comfortable", drawing quotes as in Fig. 3B, and using a sarcastic voice tone to underline it is joking.
3. Practice. Positive feedback to users who correctly complete a step at the first attempt is given in different ways based on the specific step. After the first and the third correctly completed steps, the APA sings for a few seconds a well-known Italian pop song. The two songs are different, and the original lyrics have been changed to refer to the involved part of the body (feet or hands) placed correctly. After the second step (head), the APA says "Exact" and then pretends to be surprised (with its voice and facial expression) of having a "fabulous" pair of shoes (Fig. 3C). This comment is meant to help users remember the correct position of the head, since the shoes can be seen when the head is placed in the correct position. If users do not place head, feet or hands correctly in the first attempt, the same two songs with slightly modified lyrics and a similar comment about shoes are used in telling them that the involved body part is not placed correctly. In this way, as in the second variant, users are exposed to the emotion appeal, regardless of the outcome of their first attempt. The demonstrations after the subsequent wrong attempts and the eventual acknowledgment of step completion are the same used in the first variant.
4. Final feedback. The APA tells the same sentences used in the first variant, but it adds the following jokes. When the APA says that users have made one or more errors, it adds a joke, saying that this will not make them lose points in their driving license (Italian driving laws are based on a system of penalty points removed from the driving license for violations), but it is better to try again. Before saying that users should assume the brace position faster, the APA mimics the well-known synthetic voice that announces delays in Italian train stations to say that users assumed the brace position with a considerable delay. Finally, when the APA tells users that they have assumed the correct position quickly, it smiles and nominates them "on-board Yoga master", associating the brace position to Yoga postures.

The fourth variant uses fear and humor appeals together, combining the same additions made to the second and third variants. In most cases, we were able to have the humor appeal precede the fear appeal, since this allows remarking the taught concept after the laugh as suggested in [27]. Only in three cases, we had to present the fear appeal before the humor appeal because of the content of the sentences that could not be combined in the other way. In most cases, the combined sentence is just the sequence of the two sentences with the common (if any) parts removed, and the associated body language, facial expressions, and voice tones. In the remaining two cases, the combined sentence adds a short switch between the humor and the fear appeal (e.g., the APA says "Seriously" or "Joking aside") to make it explicit that it has stopped joking, and a serious remark begins.

4 Experimental Evaluation

The study followed a 2 x 2 between-participants design: the independent variable Humor Appeal indicates if the APA appeals to humor and has two levels (Yes, No); the independent variable Fear Appeal indicates if the APA appeals to fear and has two levels (Yes, No). Each participant was assigned to one of the four groups reported in Table 1.

TABLE 1
Participants' Groups and Independent Variables

	Humor Appeal	No Humor Appeal
Fear Appeal	In this group, the APA used both humor and fear	In this group, the APA used only fear
No Fear Appeal	In this group, the APA used only humor	In this group, the APA did not use emotion appeals

We formulated the following hypotheses for the study. The first (manipulation check) states that the resort to a humor (resp. fear) appeal by the APA should actually arouse humor (resp. fear) in participants. If the manipulation check is successful, then we hypothesized that the different appeals could also change participants' mood, which is a factor that could affect performance of taught procedures [18]. Finally, we hypothesized that the humor appeal could influence liking of the application and perception of the APA in the same way it did in studies on human educators (e.g., [10], [11]). The study is instead exploratory about the effects of the fear appeal on liking and perception of the agent as well as the effects of humor and fear appeals on knowledge acquisition and transfer.

4.1 Materials

The application containing the APA was implemented using Unity 5.4, and ran on a Samsung Galaxy Tab S2 tablet with a 9.7-inch screen. Participants listened to the audio through a pair of Sennheiser HD 215 closed headphones. To collect data from participants, we employed the materials described in the following subsections.

4.1.1 Demographic questionnaire

The demographic questionnaire asked participants about their gender and age. In addition, it assessed individual

differences in frequency of air travel, by asking participants to count their number of flights in the last two years, as in [17]. Each flight had to be counted individually, so for example a round trip from airport A to airport C via a connection through airport B resulted in four flights.

4.1.2 Fear Questionnaire

Since individual phobic and anxiety levels of participants could affect their reaction to fear appeals, we administered the Fear Questionnaire [50]. The Fear Questionnaire calculates respondents' phobic and anxiety levels, respectively summing their ratings to 15 and 5 items on a 9-point scale (from 0 to 8).

4.1.3 Sense of Humor Questionnaire

Since participant's attitude towards humor could affect reaction to humor appeals, we administered the Sense of Humor Questionnaire in the 6-items form [51]. The questionnaire calculates respondents' attitude towards humor by summing their ratings to 6 items on a 4-point scale (1 to 4), after inverting the negative items.

4.1.4 Aroused fear and humor questionnaire

We administered a questionnaire, adapted from [35], which asked participants to rate their level of agreement on 12 statements using a 10-value Likert scale (0 to 9, with 9 indicating highest agreement). The first 6 statements were about aroused fear: i) "This app scared me about the dangers of not assuming the brace position", ii) "This app made me think a great deal of the dangers of not assuming the brace position", iii) "This app reminded me of how risky it is to not assume the brace position", iv) "Apps like this truly make me afraid to not assume the brace position", v) "I found myself feeling very frightened when I used this app", vi) "This is the kind of app that makes me sick to my stomach". The other 6 statements were about aroused humor: i) "I think the app I just used is very funny", ii) "I found myself laughing when I used this app", iii) "One of the things I liked about this app was how funny it was", iv) "I enjoyed the humor used in this app", v) "This app is not at all amusing", vi) "I found myself feeling very good after using this app". Aroused Fear was calculated as the average of the ratings to the first 6 statements (Cronbach's alpha = 0.87) and Aroused Humor as the average of the ratings to the other 6 statements, after inverting the scale of the negative statement (Cronbach's alpha = 0.91).

4.1.5 POMS-SF questionnaire

To measure mood, we used the POMS-SF questionnaire [52], asking participants to rate 37 sensations from 0 to 4 (0 indicates that participants did not feel that sensation and 4 that they felt it extremely). The 37 sensations are grouped to form 6 subscales: Vigor (6 items), Tension (6 items), Confusion (5 items), Anger (7 items), Depression (8 items), Fatigue (5 items). In the first subscale, higher values are more desirable because they indicate greater amount of vigor, while in the other subscales lower values are more desirable because they indicate smaller amount of negative sensations. An overall mood indicator, called Total Mood Disturbance, is calculated by inverting the first subscale and summing up all the values. The lower the value of Total Mood Disturbance, the less disturbed is the participant's mood. Since mood can vary with the participant, we measured it before (pre-test) and after (post-test) using the application, and we calculated the differences between the post-test and pre-test values.

4.1.6 Liking questionnaire

We administered a questionnaire, partially inspired by [36], which asked participants to rate their level of agreement on 7 statements using a 7-value Likert scale (1 to 7, with 7 indicating highest agreement). The statements were: i) "This app truly held my interest", ii) "This app was very boring", iii) "This is the kind of app that I'm likely to sleep through", iv) "This app is memorable", v) "I like this app very much", vi) "This app is cool", vii) "This app is appealing". After inverting negative items, Liking was calculated by averaging the ratings (Cronbach's alpha = 0.87).

4.1.7 APA perception questionnaire

We administered a 26-item questionnaire that asked participants to rate specific aspects of the APA on a 7-value scale (1 to 7, with 1 = "Not at all" and 7 = "Very"). Items were grouped to form the following subscales: Helpful (how much the APA was perceived as helpful for concentrating on and understanding the procedure), Aesthetic (how much the APA was perceived as aesthetically pleasing and attractive), Credible (how much the APA was perceived as trustworthy and expert), Sociable (how much the APA was perceived as sociable), Appropriate (how much the APA was perceived as appropriate to the situation), and Calm (how much the APA was perceived as calm).

To evaluate factorial validity of the questionnaire items, we performed exploratory factor analyses with principal component extraction and Varimax rotation. Before conducting the factor analysis, we built the correlation table among the 26 items, which revealed that one of the items ("Does it speak too fast?") did not correlate with any other item. After removal of this item, factor analysis was carried out on the 25 items listed in Table 2. Post-hoc indicators of data and sampling adequacy for the factor analysis (Kaiser-Meyer-Olkin, KMO, and Bartlett's Test of Sphericity) showed that KMO was 0.82, and Bartlett's test was significant ($p < .001$). The analysis confirmed a six-factor structure, which explained 69.6 percent of variance. Helpful, Aesthetic, Credible, Sociable, Appropriate, and Calm were calculated by averaging the answers to the items in each subscale (Cronbach's alphas were respectively 0.84, 0.82, 0.87, 0.70, 0.67, and 0.70).

4.1.8 Knowledge test

To assess acquisition of knowledge and its transfer to the real world, we asked participants to assume a brace position for an emergency landing, before (pre-test) and after (post-test) using the application. To carry out this test, we prepared two seats, one facing the back of the other as in two adjacent rows of an aircraft cabin. Participants sat on the seat in the back row and could see the back of the other seat in front of them. Experimenters checked the differences between the position assumed by participants and the taught brace position, which is the one depicted in Fig. 1. To keep track of the errors, experimenters used a

TABLE 2
APA PERCEPTION QUESTIONNAIRE

Aesthetic
Has it a realistic appearance?
Has it an attractive appearance?
Does it move in a natural way?
Is it well dressed?
Does it use facial expressions in a natural way?
Is its gaze natural?
Appropriate
Has it the appearance of a professional flight attendant?
Is its facial expression appropriate for the situation?
Does it behave in a way that is appropriate for the situation?
Calm
Is it angry, tense, nervous?
Is it serene, relaxed, calm?
Sociable
Is it cold, detached?
Is it unkind, inattentive?
Is it sociable, kind?
Is it boring?
Credible
Does it inspire trust?
Does it give a sense of confidence?
Is it intelligent?
Is it reliable?
Is it well informed?
Helpful
Does it give you clear and easy-to-understand instructions?
Does it make you pay attention to the instructions?
Does it help you to focus?
Is it useful?
Do you like to receive instructions from it?

knowledge test sheet that included a checklist with the following errors: i) feet not firmly tucked on the floor, ii) feet not placed behind the knees, iii) head not placed against the seat in front, and iv) hands not tucked behind the knees or below the tights. For each error in the checklist, experimenters observed participants and ticked the checkbox if the error was made. Experimenters were trained to rapidly recognize if participants had made the different errors by applying simple objective criteria (e.g., if the soles of the shoes were fully or partly detached from the floor, then the feet were not firmly tucked). Experimenters also annotated on the knowledge test sheet the time it took participants to assume the position. Time was measured with a chronometer application on a mobile phone.

4.2 Participants

The evaluation involved a sample of 94 participants (75 M, 19 F). They were undergraduate students in computer science as well as people from other occupations. Their age ranged from 19 to 29 (M=22.06, SD=1.77) and their number of flights in the last two years ranged from 0 to 15 (M=2.03, SD=2.84). Phobic level of participants, calculated with the Fear Questionnaire described in Section 4.1.2, ranged from 4 to 60 (M=22.89, SD=11.66), while anxiety level ranged from 0 to 29 (M=12.97, SD=6.26). Participants' attitude towards humor, calculated with the Sense of Humor Questionnaire described in Section 4.1.3, ranged from 16 to 23 (M=19.28, SD=1.73). Experimenters assigned participants to the groups trying to minimize differences between groups in the results of the demographic questionnaire, the Fear Questionnaire, and the Sense of Humor Questionnaire. Groups were also balanced in the pre-test values of Total Mood Disturbance, Vigor, Tension, Confusion, Anger, Depression, Fatigue, errors and time to assume the brace position. Each of the variables assessed before trying the experimental condition was submitted to a one-way ANOVA that confirmed the lack of significant differences between the four groups.

4.3 Measures

We measured the following dependent variables: i) Aroused Fear and Aroused Humor, with the aroused fear and humor questionnaire (Section 4.1.4); ii) Total Mood Disturbance Difference, Vigor Difference, Tension Difference, Confusion Difference, Anger Difference, Depression Difference, and Fatigue Difference, calculated as the differences between the post-test and pre-test values of the POMS-SF questionnaire described in Section 4.1.5 (a positive value for Vigor Difference and negative values for the other mood variables indicate mood improvement); iii) liking, with the liking questionnaire (Section 4.1.6); iv) Helpful, Aesthetic, Credible, Sociable, Appropriate, and Calm, with the APA perception questionnaire (Section 4.1.7); v) Error Difference and Time Difference, calculated as the differences between the post-test and pre-test values of the knowledge test described in Section 4.1.8 (negative values indicate improvement).

4.4 Procedure

The experimenter told participants that we were testing an application that teaches the position to be assumed during an aircraft emergency landing. Consent for participation and for recording verbal answers was obtained. Participants were also informed that they could refrain from continuing the experiment at any time, without providing a reason to the experimenter. Then, the experimenter tested participants' initial knowledge about the brace position. In particular, participants sat on the seat described in Section 4.1.8, and the experimenter told them to imagine to be on an aircraft after being warned to prepare for impact. Then, participants were asked to physically assume the position they would assume in that circumstance, and to do it as fast as possible after a start command from the experimenter. They had to inform the experimenter when finished by saying "stop". The experimenter kept track of the errors in the assumed position with the knowledge test sheet (Section 4.1.8), and annotated the time spent to assume the position, from start to stop command. Then, participants filled the demographic questionnaire, the Fear Questionnaire, the Sense of Humor Questionnaire, and the POMS-SF questionnaire. The experimenter gave participants the tablet, and helped them wear a pair of closed earphones. They tried the condition assigned to them, going through all the phases of the lesson described in Section 3. Then, they filled the POMS-SF questionnaire, the liking questionnaire, the APA perception questionnaire, and the aroused fear and humor questionnaire. After that, the experimenter tested again participants' knowledge about the brace position, and annotated time and errors. Finally, participants' qualitative feedback was acquired by means of a

brief semi-structured interview. Participants were asked about the aspects they liked and disliked, and about the things they would change in the application or in the APA.

5. Results

Since our groups differed on two independent variables as described in Table 1, we analyzed the results using a between-subjects two-way ANOVA for each of the dependent variables described in Section 4.3. In case of interaction, we analyzed each simple effect using a between-subjects one-way ANOVA with the mean square error from the initial analysis as error term. Effect sizes are reported as partial eta squared (η_p^2).

5.1 Aroused Fear

For Aroused Fear (Fig. 4A), the analysis revealed a main effect of Fear Appeal, $F(1,90)=56.24$, $p<0.001$, $\eta_p^2=0.39$, no main effect of Humor Appeal, and no interaction. When the APA used fear, Aroused Fear for participants was higher ($M=3.95$, $SD=1.61$) than when it did not ($M=1.48$, $SD=1.58$).

5.2 Aroused Humor

For Aroused Humor (Fig. 4B), the analysis revealed a main effect of Humor Appeal, $F(1,90)=70.44$, $p< 0.001$, $\eta_p^2=0.44$, no main effect of Fear Appeal, and an interaction between Fear Appeal and Humor Appeal, $F(1,90)=3.95$, $p<0.05$, $\eta_p^2=0.04$. When the APA used humor, Aroused Humor was higher ($M=5.35$, $SD=2.10$) than when it did not ($M=2.24$, $SD=1.52$). To investigate the interaction, we analyzed simple effects of Humor Appeal separately at the two levels of Fear Appeal, and simple effects of Fear Appeal separately at the two levels of Humor Appeal. At both levels of Fear Appeal, Aroused Humor was higher when the APA used humor (without fear appeal: $M=5.93$, $SD=1.97$; with fear appeal: $M=4.81$, $SD=2.12$) than when it did not (with no fear appeal: $M=2.06$, $SD=1.81$; with fear appeal: $M=2.42$, $SD=1.15$), $p<0.001$ for both. With no humor appeal, we found no statistically significant difference in Aroused Humor between participants whose APA used fear and participants whose APA did not. Instead, with humor appeal, Aroused Humor was lower when the APA used fear ($M=4.81$, $SD=2.12$) than when it did not ($M=5.93$, $SD=1.97$), $p<0.05$.

5.3 Mood

For Total Mood Disturbance Difference (Fig. 5A), the analysis revealed a main effect of Fear Appeal, $F(1,90)=9.58$, $p<0.01$, $\eta_p^2=0.10$, a main effect of Humor Appeal, $F(1,90)=6.66$, $p<0.05$, $\eta_p^2=0.07$, and no interaction. When the APA used fear, the value of Total Mood Disturbance Difference was higher ($M=-0.34$, $SD=4.26$) than when it did not ($M=-3.53$, $SD=6.02$). Instead, when the APA used humor, Total Mood Disturbance Difference was lower ($M=-3.26$, $SD=5.31$) than when it did not ($M=-0.62$, $SD=5.28$). Fig. 7 shows the means of Total Mood Disturbance Difference for each value of the independent variables. Considering the subscales of mood, the analysis revealed the following main effects. When the APA used fear, Vigor Difference (Fig. 5B) was lower ($M=-1.83$, $SD=2.48$) than when it did not ($M=-0.40$, $SD=3.22$), $F(1,90)=6.19$, $p<0.05$, $\eta_p^2=0.06$. When the APA used humor, Vigor Difference was higher ($M=-0.55$, $SD=2.76$) than when it did not ($M=-1.68$, $SD=3.05$), $F(1,90)=3.95$, $p<0.05$, $\eta_p^2=0.04$. When the APA used fear, Confusion Difference (Fig. 5C) was higher ($M=-0.43$, $SD=1.54$) than when it did not ($M=-1.13$, $SD=1.50$), $F(1,90)=5.13$, $p<0.05$, $\eta_p^2=0.05$.

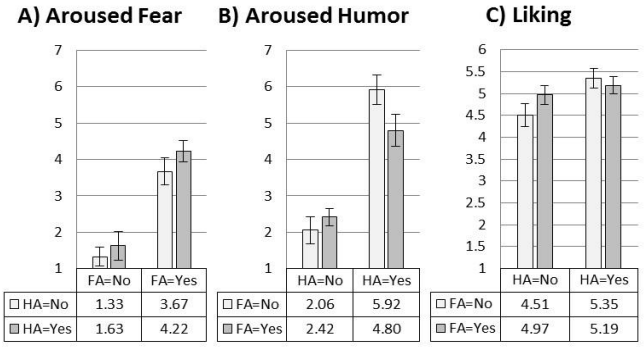

Fig. 4. Means of A) Aroused Fear, B) Aroused Humor, and C) Liking. HA = Humor Appeal, FA = Fear Appeal. Capped vertical bars indicate ± SE.

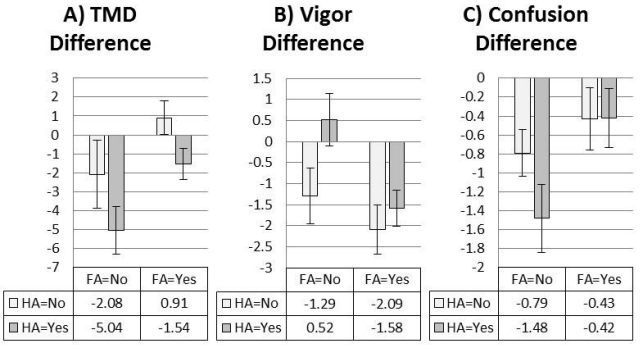

Fig. 5. Means of A) Total Mood Disturbance (TMD) Difference, B) Vigor Difference, and C) Confusion Difference. HA = Humor Appeal, FA = Fear Appeal. Capped vertical bars indicate ± SE.

5.4 Liking

The analysis revealed a main effect of Humor Appeal on Liking (Fig. 4C), $F(1,90)=5.59$, $p<0.05$, $\eta_p^2=0.06$, no main effect of Fear Appeal, and no interaction. When the APA used humor, Liking was higher ($M=5.27$, $SD=0.98$) than when it did not ($M=4.73$, $SD=1.18$).

5.5 Perception of the APA

The analysis revealed no main effect of Fear Appeal, no main effect of Humor Appeal, and no interaction, on Helpful, Aesthetic, and Credible. A main effect of Humor Appeal, $F(1,90)=79.87$, $p<0.001$, $\eta_p^2=0.47$, no main effect of Fear Appeal, and no interaction were found for Sociable (Fig. 6A). When the APA used humor, Sociable was higher ($M=6.36$, $SD=0.72$) than when it did not ($M=4.76$, $SD=0.98$). A main effect of Humor Appeal, $F(1,90)=5.17$, $p<0.05$, $\eta_p^2=0.05$, no main effect of Fear Appeal, and no interaction were found also for Appropriate (Fig. 6B). However, in this case, Appropriate was lower when the APA used humor ($M=4.52$, $SD=1.18$) than when it did not ($M=5.08$, $SD=1.21$). Finally, considering Calm (Fig. 6C), the analysis revealed a main effect of Humor Appeal, $F(1,90)=13.85$, $p<0.001$,

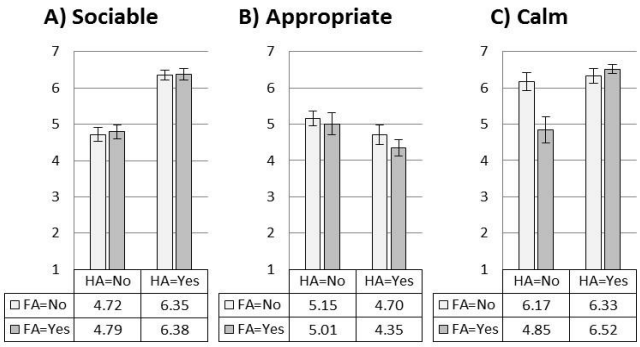

Fig. 6. Means of A) Sociable, B) Appropriate, and C) Calm. HA = Humor Appeal, FA = Fear Appeal. Capped vertical bars indicate ± SE.

η_p^2=0.13, a main effect of Fear Appeal, F(1,90)=5.21, p<0.05, η_p^2=0.06, and interaction, F(1,90)=9.45, p<0.01, η_p^2=0.10. When the APA used humor, Calm was higher (M=6.43, SD=0.79) than when it did not (M=5.52, SD=1.62). When the APA used fear, Calm was lower (M=5.70, SD=1.54) than when it did not (M=6.24, SD=1.06). To investigate the interaction, we analyzed simple effects of Humor Appeal separately at the two levels of Fear Appeal, and simple effects of Fear Appeal separately at the two levels of Humor Appeal. With no fear appeal, we found no statistically significant difference in Calm between the APA using humor or not. Instead, with fear appeal, Calm was higher when the APA used humor (M=6.52, SD=0.63) than when it did not (M=4.85, SD=1.75), p<0.001. With no humor appeal, Calm was lower when the APA used fear (M=4.85, SD=1.75) than when it did not (M=6.17, SD=1.19), p<0.001. Instead, with humor appeal, we found no statistically significant difference in Calm between the APA using fear or not.

5.6 Knowledge transfer

The analysis found no main effect of Fear Appeal, no main effect of Humor Appeal, and no interaction, in Error Difference as well as Time Difference. All four groups improved their knowledge as we will discuss in the next session.

6. Discussion

Results of the manipulation check showed that the use of humor and fear appeals by the APA was effective in arousing the intended emotions. Regarding fear, the Aroused Fear score was significantly higher in participants exposed to a fear appeal by the APA. Participants in both groups that were not exposed to the fear appeal had low scores of Aroused Fear (on a 0 to 9 scale, means were 1.33 without emotion appeals, and 1.63 with only humor appeal). On the contrary, the fear appeal aroused a larger level of fear in the other two groups (means were, respectively, 3.67 in the group exposed only to fear appeal, and 4.22 in the group exposed to both fear and humor appeals). The level of fear aroused by the APA through the fear appeal was moderate, which is a positive result, because a low level of fear is unlikely to motivate people to consider how to cope with a risk, while too much fear could induce defensive responses which are detrimental to the effectiveness of the communication [53].

Regarding humor, the Aroused Humor score was significantly higher in participants exposed to the humor appeal. Moreover, we found an interaction between Fear Appeal and Humor Appeal for Aroused Humor. The analysis of simple effects showed that combining fear appeal with humor appeal led to a lower Aroused Humor with respect of using only humor appeal. When the APA described the severe consequences of not assuming the brace position, the application was inevitably perceived as less humorous.

In addition to arousing the intended emotions, when the APA used emotion appeals, it significantly affected participants' mood, confirming our hypothesis. Indeed, when the APA used the humor appeal, Total Mood Disturbance decreased significantly more than when the APA did not use it. On the contrary, when the APA used the fear appeal, Total Mood Disturbance decreased less (when fear appeal was combined with humor appeal) or even increased (when fear appeal was used alone), as shown in Fig. 5A. Since lower values of Total Mood Disturbance indicate a better mood, this result suggests that the use of humor appeal improved participants' mood, while the use of fear appeal had a negative effect on it. This is particularly important because people's mood can affect their performance [54], [55]. More specifically, in the domain of aviation safety education, Tehrani and Molesworth [18] found that participants exposed to a positive mood manipulation performed an evacuation exercise better than participants who were exposed to a negative mood manipulation. This suggests that the use of humor by an APA for in-flight safety briefings might provide an advantage in terms of passengers' mood as well as performance.

Some additional insights come from the analysis of the subscales of mood. When the APA used fear, the Vigor score decreased significantly more than when the APA did not use it. Participants who were exposed to the fear appeal might have felt more threatened, and might have become less lively, cheerful, and energetic, which are three of the adjectives in the Vigor subscale. Interestingly, as shown in Fig. 5B, Vigor decreased also when the APA used no emotion appeals, while it increased only when the APA used humor without fear. This contributed to the significant result we found for the use of humor appeal on Vigor Difference, but also suggests that other aspects of the application might play a role in decreasing Vigor, and that this negative effect is mitigated by the use of humor.

Confusion decreased in all four groups. This may suggest that the APA was able to explain the procedure clearly, and focus participants' attention. It is interesting to note that while some studies concerning humor in classrooms (e.g., [26]) found that students were confused by humor, we did not find negative effects of the APA's humor on confusion. The humor appeal used by the APA did not distract participants from the taught content because the humor was related to the content, as suggested by [27]. The improvement in confusion was instead significantly lower when the APA used fear than when it did not. Aroused fear might have limited the positive effect of the APA on

confusion.

It is important to note that mood and its subscales were self-reported by participants. Although the POMS-SF questionnaire we employed has been extensively used and validated [52], [56], self-reported measures might be biased by different factors such as social desirability and participants' inability to report their own mood. Therefore, further studies employing more than one type of mood measure should be conducted to better support our results about mood.

Considering our hypothesis about liking of the application, only the humor appeal had a statistically significant effect. The two groups in which the APA used humor liked the application more than the other two groups. This positive effect is consistent with the studies on human educators, which showed that humor can increase educator's likeability [10]. Fig. 4C might suggest that the use of fear appeal increased liking with respect to the condition with no appeals, and that it reduced liking with respect to the condition with only humor. However, we found no significant main effect of fear appeal on liking and no interaction between humor and fear appeal on this aspect, so additional research is needed to assess whether fear appeals could actually affect liking.

As hypothesized, the use of emotion appeals affected participants' perception of the APA. The effect concerned some of the subscales. More precisely, we did not find statistically significant differences on how much the participants perceived the APA as helpful. Participants in all groups positively rated this aspect (an average of 5.64 on a 1 to 7 scale). Despite the fact that the Aesthetic score was slightly higher with humor appeal, we did not find any main effect nor interaction. Similarly, we did not find statistically significant differences for the credibility of the APA, which was high in all four groups (an average of 5.29 on a 1 to 7 scale), including those exposed to humor appeal. This shows that participants trusted the APA and perceived its expertise regardless of the appeals used. This result may suggest that the amount of humor used by the APA was not excessive, consistently with the fact that too much humor could decrease credibility [22].

Results show very clearly that the APA was perceived as more sociable when it used humor. No significant effect on sociability was found for fear appeal, while the improvement using humor appeal was considerable (+1.61 on a 1 to 7 scale). The only negative effect of humor appeal was found for APA's appropriateness, which was lower (-0.56 on a 1 to 7 scale) when the APA used humor. The semi-structured interview provided insights on this. A few (three) participants said that the use of humor was not appropriate because the topic or the situation were serious, while some (nine) said that the APA used too much humor. Interestingly, five of these twelve participants as well as other five participants specifically identified the singing of songs by the APA as an aspect they found inappropriate or they did not like. This might suggest that the negative effect on appropriateness was due mainly to this humor element, and further studies with and without humorous songs are needed to clarify their role.

Emotion appeals affected participants' perception of the APA as calm. While, in general, the fear appeal made participants perceive the APA as more tense, and the humor appeal made them perceive it as more calm, the significant interaction provided an interesting insight: as suggested by Fig. 6C, humor had almost no effect on perceived calmness when fear was not used, while it counteracted the effect of fear when it was used.

Finally, we found no significant effect of fear and humor appeals, and no significant interaction, on knowledge transfer. Therefore, we considered all the groups together and compared pre-test with post-test knowledge in a repeated measures design to assess if the application was effective to decrease errors in assuming the brace position and the time employed. Paired samples t-tests indicated that differences between post-test and pre-test values in both errors and time were statistically significant, $p<0.001$ for both variables. Users made an average of 2.70 (SD=0.84) errors before using the application and only 0.21 (SD=0.46) after using it (Fig. 7A), and time to assume the brace position decreased from 4.54 (SD=3.08) to 3.03 (SD=1.71) seconds (Fig. 7B). It is important to remark that all participants made errors in performing the procedure before using the application, while 80.9% of participants made no errors after using it. Moreover, the time to assume the brace position decreased by about one third. Additional paired samples t-tests carried out on each group independently showed that error decrease was significant in each group, $p<0.001$. Therefore, the application had positive learning effects regardless of the emotion appeals used by the APA. However, since the maximum number of possible errors was four, future studies on more complex procedures, made of a larger number of steps, could be carried out to further assess the role of emotion appeals on learning. In addition, it must be noted that participants used the application in a laboratory setting, and wore closed headphones. In this way, they could use the application without being distracted. This context can be different from more naturalistic settings (e.g., airport or aircraft cabin). In a setting where users could be distracted by external factors, emotions aroused by humor and fear appeals might result in increased attention and have an impact on knowledge retention [57], [58]. Therefore, it would be interesting to replicate the experiment outside the lab.

We carried out additional analyses on participants' interaction data logged by the application, i.e. participants' acceptance of the APA invitation to repeat the trial, number of errors made during the practice phase, time taken to place the feet, the head, and the hands in that phase. Almost all participants accepted the APA invitation to repeat the trial, and the differences between groups did not reach statistical significance. We did not find significant differences between groups for errors and times in the first trial, and for errors in the second trial. Differences in the times in the second trial reached significance: times were larger in the conditions with the emotion appeals. However, it must be noted that times in the second trial were generally lower than those in the first trial for all groups, reflecting the improvement in learning the procedure, and the larger times found with emotion appeals could be simply due to the longer sentences used for feedback in those conditions.

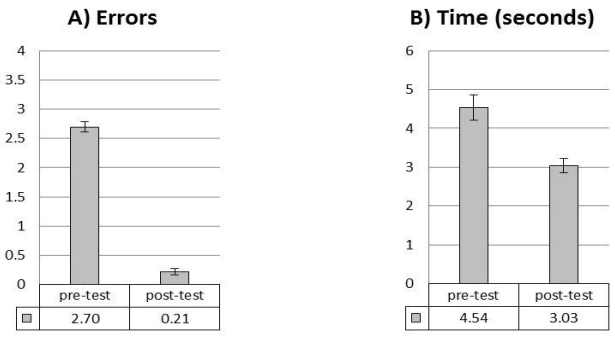

Fig. 7. Means of A) errors and B) time (in seconds) to assume the brace position before and after using the application. Capped vertical bars indicate ± SE.

In summary, our study shows that several results found for humor and fear appeals used by human educators transfer to APAs, and extends the investigation of effects to mood. More precisely, we have shown that an APA using a humor appeal can i) improve people's mood, which could bring performance improvements, ii) increase liking of the application, iii) appear more sociable, but also iv) appear less appropriate for the situation. An APA using a fear appeal can i) arose a moderate level of fear, which should motivate people to cope with the presented risks, but also ii) worsen mood.

The qualitative feedback obtained with the interviews helped us to better understand which types of humor the APA should use. Among the 47 participants who listened to the APA using the humor appeal, 30 explicitly said that they liked the use of humor. However, only nine said that they liked the use of the humorous songs, and ten explicitly said that they did not like the songs. Only one participant said that he did not like the jokes, while eleven participants mentioned jokes or cited one of them when they positively commented about humor. Interestingly, nine participants who listened to the APA not using humor (six exposed to fear appeal and three not exposed to emotion appeals) said that it was too serious. Overall, the qualitative feedback suggested that participants generally welcomed the use of humor by the APA, and that jokes were more appreciated and less criticized than humorous songs.

Among the 47 participants who listened to the APA using the fear appeal, 16 said that they appreciated the explanation of the negative consequences of wrong actions, which were part of the fear appeal, and no participants made negative comments about them. Six participants exposed to the fear appeal (and not to the humor appeal) said that the APA was too serious. Interestingly, no participant who listened to the APA that used both humor and fear made similar comments. Interestingly, ten participants (3 exposed and 7 not exposed to the fear appeal) suggested to visually show the negative consequences of wrong actions (e.g., with an illustration of the physical injuries). Overall, the qualitative feedback suggested that showing and explaining the consequences of wrong actions could be appreciated by users, and the combination of the fear and humor appeals might be useful to explain the negative consequences of wrong actions without making users negatively perceive the APA as too serious.

It is important to note that the study was focused on a safety education domain, so one cannot assume that the findings generalize to all the different domains in which APAs are used. For example, the use of humor appeals might be perceived as more appropriate if the domain does not involve serious risks, and at the same time, the use of fear appeals may not be obvious in those domains. However, humor depends on culture and age, so the results could vary if different populations are considered. Moreover, our sample included participants with an attitude towards humor ranging from 16 to 23 (M=19.28, SD=1.73) in a 4 to 24 scale, so involving participants with less sense of humor would be interesting to assess whether the results could be extended or not. Future studies may also assess the possible role of participants' individual differences (e.g., different sense of humor, or previous level of fear of the considered risk).

Within the same population, the findings in our study may be probably extended to other health and safety education domains beyond aviation, since health and safety education typically needs to make people aware of risks (e.g., the risk of illness if proper hygiene procedures or lifestyle are not followed, the risk of injuries if safe driving rules are not followed,…), and teach the correct procedures to cope with them (e.g., how to protect skin from ultraviolet radiation exposure, how to wear safety restraints,…).

7. Conclusions and Future Work

This paper explored the use of humor and fear appeals by an APA, and is the first to test their combined use and to assess the possible interactions following a 2 x 2 factorial design. The APA we developed taught an aviation safety procedure. We carried out an evaluation with 94 participants, showing that the APA was successful in teaching the procedure regardless of the appeals used. However, resort to humor and fear appeals by the APA significantly affected, and in different ways, participants' mood, liking of the application, and perception of the APA.

As discussed in the previous section, further studies employing more than one type of mood measure should be conducted to better support our findings about mood. The study was focused on the domain of aviation safety education, and findings might be extended to other health and safety education domains that share similar needs. Since the APA is programmable, it can be used also in domains unrelated to health and safety. This would allow to investigate possible changes in the effects, e.g. testing if the positive effects of humor appeals (e.g., on liking and perception of the APA as sociable) may be even more evident or their negative effect (on perception of the APA as appropriate) may be reduced.

Since the effects of humor appeals may vary with age and culture, there is a need to investigate them more broadly on different populations. A possible way to achieve this purpose would be to use remote evaluation: the APA might be programmed to use different appeals in a public mobile app, which could include a questionnaire to remotely assess opinions of users of different ages and cultures. The questionnaire could also be used to assess

different aspects of users' profiles (e.g., sense of humor and fear of flying) to study if such individual differences could play a role on the effects of humor and fear appeals.

In addition, further research is needed to better understand the effect of humor and fear appeals on learning, conducting studies with more complex procedures that require to remember a larger number of steps and details than the brace procedure. Finally, since humor and fear appeals can arouse emotions, and emotions can play a role in knowledge retention [57], [58], a further study could test knowledge retention over time after exposure to different emotion appeals.

Acknowledgment

This work was supported in part by a grant from US Federal Aviation Administration (FAA). We are grateful to David Weed (CAMI, FAA) for his feedback and encouragement. The authors wish to thank Nicola Zangrando (HCI Lab, University of Udine), who carried out 3D modeling and animation activities for the development of the application.

References

[1] W. L. Johnson and J. W. Rickel, "Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments," *Int. J. Artif. Intell. Educ.*, vol. 11, pp. 47–78, 2000.

[2] J. Cassell and K. R. Thorisson, "The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents," *Appl. Artif. Intell.*, vol. 13, no. 4–5, pp. 519–538, May 1999.

[3] H. Maldonado et al., "We Learn Better Together: Enhancing eLearning with Emotional Characters," in *Proc. 2005 Conf. Computer Support for Collaborative Learning 2005: The Next 10 Years!*, 2005, pp. 408–417.

[4] A. L. Baylor and S. Kim, "Designing nonverbal communication for pedagogical agents: When less is more," *Comput. Human Behav.*, vol. 25, no. 2, pp. 450–457, 2009.

[5] S. D'Mello, R. W. Picard, and A. C. Graesser, "Toward an Affect-Sensitive AutoTutor," *IEEE Intell. Syst.*, vol. 22, no. 4, pp. 53–61, Jul. 2007.

[6] A. Sarrafzadeh, S. Alexander, F. Dadgostar, C. Fan, and A. Bigdeli, "'How do you know that I don't understand?' A look at the future of intelligent tutoring systems," *Comput. Human Behav.*, vol. 24, no. 4, pp. 1342–1363, 2008.

[7] X. Mao and Z. Li, "Agent based affective tutoring systems: A pilot study," *Comput. Educ.*, vol. 55, no. 1, pp. 202–208, 2010.

[8] D. Zillmann, B. R. Williams, J. Bryant, K. R. Boynton, and M. A. Wolf, "Acquisition of information from educational television programs as a function of differently paced humorous inserts.," *J. Educ. Psychol.*, vol. 72, no. 2, pp. 170–180, Apr. 1980.

[9] N. Kher, S. Molstad, and R. Donahue, "Using humor in the college classroom to enhance teaching effectiveness in 'dread courses,'" *Coll. Stud. J.*, vol. 33, no. 3, pp. 400–406, 1999.

[10] C. Conkell, C. Imwold, and T. Ratliffe, "The effects of humor on communicating fitness concepts to high school students," *Phys. Educ.*, vol. 56, pp. 8–18, 1999.

[11] M. Houser, R. Cowan, and D. West, "Investigating a New Education Frontier: Instructor Communication Behavior in CD-ROM Texts - Do Traditionally Positive Behaviors Translate into This New Environment?," *Commun. Q.*, vol. 55, no. 1, pp. 19–38, 2007.

[12] R. W. Rogers and C. W. Deckner, "Effects of fear appeals and physiological arousal upon emotion, attitudes, and cigarette smoking.," *J. Pers. Soc. Psychol.*, vol. 32, no. 2, pp. 222–230, 1975.

[13] K. Witte and M. Allen, "A meta-analysis of fear appeals: implications for effective public health campaigns," *Heal. Educ. Behav.*, vol. 27, no. 5, pp. 591–615, Oct. 2000.

[14] A. Mukherjee and L. Dubé, "Mixing emotions: The use of humor in fear advertising," *J. Consum. Behav.*, vol. 11, pp. 147–161, 2012.

[15] N. D. Weinstein, "Accuracy of smokers' risk perceptions," *Ann. Behav. Med.*, vol. 20, no. 2, pp. 135–140, Jun. 1998.

[16] H. A. Deery, "Hazard and Risk Perception among Young Novice Drivers," *J. Safety Res.*, vol. 30, no. 4, pp. 225–236, Dec. 1999.

[17] C. L. Corbett, G. A. McLean, and D. K. Cosper, "Effective Presentation Media for Passenger Safety I: Comprehension of Briefing Card Pictorials and Pictograms," Technical Report DOT/FAA/AM-08/20, Washington, DC, USA, Sep. 2008.

[18] M. Tehrani and B. R. C. Molesworth, "The Effect of Mood on Performance in a Nonnormal Situation," *Aviat. Psychol. Appl. Hum. Factors*, vol. 6, no. 1, pp. 1–11, May 2016.

[19] J. A. Banas, N. Dunbar, D. Rodriguez, and S.-J. Liu, "A Review of Humor in Educational Settings: Four Decades of Research," *Commun. Educ.*, vol. 60, no. 1, pp. 115–144, Jan. 2011.

[20] R. A. Martin, *The psychology of humor: An integrative approach*. Oxford, UK: Elsevier Academic Press, 2007.

[21] S. Booth-Butterfield and M. Booth-Butterfield, "Individual differences in the communication of humorous messages," *South. Commun. J.*, vol. 56, no. 3, pp. 205–218, 1991.

[22] S. E. Torok, R. F. McMorris, and W.-C. Lin, "Is Humor an Appreciated Teaching Tool? Perceptions of Professors' Teaching Styles and Use of Humor," *Coll. Teach.*, vol. 52, no. 1, pp. 14–20, 2004.

[23] M. B. Wanzer, A. Bainbridge Frymier, A. M. Wojtaszczyk, and T. Smith, "Appropriate and Inappropriate Uses of Humor by Teachers," *Commun. Educ.*, vol. 55, no. 2, pp. 178–196, Apr. 2006.

[24] A. B. Frymier, M. B. Wanzer, and A. M. Wojtaszczyk, "Assessing Students' Perceptions of Inappropriate and Appropriate Teacher Humor," *Commun. Educ.*, vol. 57, no. 2, pp. 266–288, Apr. 2008.

[25] M. B. Wanzer, A. B. Frymier, and J. Irwin, "An Explanation of the Relationship between Instructor Humor and Student Learning: Instructional Humor Processing Theory," *Commun. Educ.*, vol. 59, no. 1, pp. 1–18, Jan. 2010.

[26] J. Bryant and D. Zillmann, "Using Humor to Promote Learning in the Classroom," in *Humor and children's development: A guide to practical applications*, P. McGhee and M. Frank, Eds. New York, NY, USA: Haworth Press, 1989, pp. 49–78.

[27] A. Ziv, "Teaching and Learning with Humor," *J. Exp. Educ.*, vol. 57, no. 1, pp. 4–15, Sep. 1988.

[28] M. Tehrani and B. R. C. Molesworth, "Pre-flight safety briefings, mood and information retention," *Appl. Ergon.*, vol. 51, pp. 377–382, 2015.

[29] 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. Stud.*, in press, 2019, doi: 10.1016/j.ijhcs.2018.07.006.

[30] B. Çelik and K. Gündoğdu, "The effect of using humor and concept cartoons in high school ICT lesson on students' achievement, retention, attitude and anxiety," *Comput. Educ.*, vol. 103, pp. 144–157, Dec. 2016.

[31] R. A. C. Ruiter, C. Abraham, and G. Kok, "Scary warnings and rational precautions: A review of the psychology of fear appeals," *Psychol. Health*, vol. 16, no. 6, pp. 613–630, Nov. 2001.

[32] K. Witte, "Putting the fear back into fear appeals: The extended parallel process model," *Commun. Monogr.*, vol. 59, no. 4, pp.

329–349, Dec. 1992.

[33] P. Kim, P. Sorcar, S. Um, H. Chung, and Y. S. Lee, "Effects of episodic variations in web-based avian influenza education: influence of fear and humor on perception, comprehension, retention and behavior," *Health Educ. Res.*, vol. 24, no. 3, pp. 369–380, Jun. 2009.

[34] J. C. Mowen, E. G. Harris, and S. A. Bone, "Personality traits and fear response to print advertisements: Theory and an empirical study," *Psychol. Mark.*, vol. 21, no. 11, pp. 927–943, 2004.

[35] B. M. J. Lee and M. A. Ferguson, "Effects of Anti-Tobacco Advertisements Based on Risk-Taking Tendencies: Realistic Fear vs. Vulgar Humor," *Journal. Mass Commun. Q.*, vol. 79, no. 4, pp. 945–963, 2002.

[36] M. J. Lee and M. Shin, "Fear Versus Humor: The Impact of Sensation Seeking on Physiological, Cognitive, and Emotional Responses to Antialcohol Abuse Messages," *J. Psychol. Interdiscip. Appl.*, vol. 145, no. 2, pp. 73–92, 2011.

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

[38] L. Chittaro, F. Buttussi, and N. Zangrando, "Desktop virtual reality for emergency preparedness: User Evaluation of an Aircraft Ditching Experience under Different Fear Arousal Conditions," in *Proc. 20th ACM Symp. Virt. Real. Soft. Tech.*, 2014, pp. 141–150.

[39] 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.

[40] R. W. Rogers, "Cognitive and physiological processes in fear appeals and attitude change: A revised theory of Protection Motivation," in *Social Psychophysiology: A sourcebook*, J. T. Cacioppo and R. E. Petty, Eds. New York, NY, USA: Guilford Press, 1983, pp. 153–176.

[41] J. Cassell, "Embodied conversational interface agents," *Commun. ACM*, vol. 43, no. 4, pp. 70–78, Apr. 2000.

[42] G. D. Chen, J. H. Lee, C. Y. Wang, P. Y. Chao, L. Y. Li, and T. Y. Lee, "An empathic avatar in a computer-aided learning program to encourage and persuade learners," *Educ. Technol. Soc.*, vol. 15, no. 2, pp. 62–72, 2012.

[43] R. Beale and C. Creed, "Affective interaction: How emotional agents affect users," *Int. J. Hum. Comput. Stud.*, vol. 67, no. 9, pp. 755–776, 2009.

[44] W. Burleson and R. W. Picard, "Gender-Specific Approaches to Developing Emotionally Intelligent Learning Companions," *IEEE Intell. Syst.*, vol. 22, no. 4, pp. 62–69, Jul. 2007.

[45] C. Okonkwo and J. Vassileva, "Affective pedagogical agents and user persuasion," in *Proc. 1st Intl. Conf. on Universal Access in Human-Computer Interaction*, 2001, pp. 397–401.

[46] B. A. Stone and J. C. Lester, "Dynamically Sequencing an Animated Pedagogical Agent," *Proc. 13th Natl. Conf. Artif. Intell.*, pp. 424–431, 1996.

[47] E. André, T. Rist, and J. Muller, "Employing AI Methods to Control the Behavior of Animated Interface Agents," *Appl. Artif. Intell.*, vol. 13, no. 4–5, pp. 415–448, 1999.

[48] H. C. Lane et al., "The Effects of a Pedagogical Agent for Informal Science Education on Learner Behaviors and Self-efficacy," in *Proc. Intl. Conf. Artificial Intelligence in Education*, 2013, pp. 309–318.

[49] C. Wang, S. Ke, H. Chuang, H. Tseng, and G. Chen, "E-learning system design with humor and empathy interaction by virtual human to improve students' learning," in *Proc. 18th Intl. Conf. on Computers in Education*, 2010, pp. 615–622.

[50] I. M. Marks and A. M. Mathews, "Brief standard self-rating for phobic patients," *Behav. Res. Ther.*, vol. 17, no. 3, pp. 263–267, 1979.

[51] S. Svebak, "The development of the Sense of Humor Questionnaire: From SHQ to SHQ-6," *Humor*, vol. 9, no. 3/4, pp. 341–361, 1996.

[52] S. Shacham, "A Shortened Version of the Profile of Mood States," *J. Pers. Assess.*, vol. 47, no. 3, pp. 305–306, 1983.

[53] P. A. Rippetoe and R. W. Rogers, "Effects of components of protection-motivation theory on adaptive and maladaptive coping with a health threat," *J. Pers. Soc. Psychol.*, vol. 52, no. 3, pp. 596–604, 1987.

[54] D. Albarracin and W. Hart, "Positive mood + action = negative mood + inaction: Effects of general action and inaction concepts on decisions and performance as a function of affect," *Emotion*, vol. 11, no. 4, pp. 951–957, 2011.

[55] R. S. Friedman, J. Förster, and M. Denzler, "Interactive Effects of Mood and Task Framing on Creative Generation," *Creat. Res. J.*, vol. 19, no. 2–3, pp. 141–162, Jul. 2007.

[56] S. L. Curran, M. A. Andrykowski, and J. L. Studts, "Short Form of the Profile of Mood States (POMS-SF): Psychometric information.," *Psychol. Assess.*, vol. 7, no. 1, pp. 80–83, 1995.

[57] E. A. Kensinger, "Remembering the Details: Effects of Emotion," *Emot. Rev.*, vol. 1, no. 2, pp. 99–113, Jan. 2009.

[58] B. Finn and H. L. Roediger, "Enhancing retention through reconsolidation: negative emotional arousal following retrieval enhances later recall," *Psychol. Sci.*, vol. 22, no. 6, pp. 781–6, Jun. 2011.

Fabio Buttussi received his PhD in computer science from the University of Udine. He is a postdoctoral research fellow at Human-Computer Interaction (HCI) Lab in the Department of Mathematics, Computer Science, and Physics of the University of Udine, Italy. His research interests are in HCI, virtual reality, artificial intelligence, and their application in health and safety.

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 is an ACM Distinguished Speaker. His major research interests are in virtual reality, mobile HCI, persuasive technology, and their applications in health and safety. He has received research grants from a wide range of organizations, including the US Federal Aviation Administration (FAA), the European Union (EU), the Italian Ministry of University and Research (MIUR), and companies such as the Benetton Group and the Intesa Sanpaolo bank group.