Are Aggressive Agents as Scary as Aggressive Humans?

Romy Blankendaal[1], Tibor Bosse[1], Charlotte Gerritsen[2], Tessa de Jong[1] and Jeroen de Man[1]
[1] VU University Amsterdam, Agent Systems Research Group
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
[2] Netherlands Institute for the Study of Crime and Law Enforcement
De Boelelaan 1077a, 1081 HV Amsterdam, The Netherlands
romy.blankendaal@gmail.com, t.bosse@vu.nl,
cgerritsen@nscr.nl, tessadejong71@msn.com, j.de.man@vu.nl

ABSTRACT
Most research into intelligent virtual agents focuses on agents with a positive stance towards the user. Nevertheless, the development of virtual agents that show aggressive behavior may also be interesting for a range of application domains, varying from aggression de-escalation training to anti-bullying education. However, ensuring that such aggressive agents achieve the desired effect is not easy, as they need to be believable in a number of aspects. In particular, they need to bring their human conversation partners into a serious state of anxiety. To investigate to what extent this can be achieved using state-of-the-art virtual agent technology, an experiment was performed in which the impact of an aggressive virtual agent was compared with that of an aggressive human. By randomly distributing a group of 28 participants over two conditions (virtual and human) and measuring their physiological and subjective emotional state before and after an aggressive outburst of their conversation partner, the difference between virtual and human aggression was studied. The results point out that both types of aggression induced a substantial stress response, but that the impact of the human aggression was higher than that of the virtual aggression.

Categories and Subject Descriptors
[**Human-centered computing**]: Empirical studies in HCI.

General Terms
Experimentation, Human Factors.

Keywords
Human-agent interaction; Aggression; Anxiety

1. INTRODUCTION
"You are about to finish your daily surveillance route, when you notice a middle-aged man with a baseball cap, who seems to be harassing a young lady. You politely ask the couple whether everything is OK. Suddenly, the man starts behaving aggressively towards you: he approaches you with clenched fists, shouting at you to mind your own business. The angry look on his face gives you the shivers. As he moves closer, you literally feel your blood running cold. Just when you are considering drawing your weapon, you realize he is just a character in a Virtual Reality environment. With a sigh of relief, you switch off your head-mounted display."

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May, 4–8, 2015, Istanbul, Turkey.*

Although this example sounds futuristic, it illustrates the impact that intelligent virtual agents (IVAs) may have on the mental and physical state of the human beings that interact with them. IVAs can be defined as 'interactive characters that exhibit human-like qualities and communicate with humans or with each other using natural human modalities such as facial expressions, speech and gesture' [2]. Recently, IVAs have become widely used for numerous applications, in particular within serious games [27]. They have been successfully applied in a variety of games, in domains ranging from military missions [33] to negotiation [20]. In such applications, IVAs play various roles in which they interact with users (mostly via conversations), for instance as a coach, instructor, mentor, therapist, or teammate [26].

Despite these developments, we are only scratching the surface of potential applications. Interestingly, the majority of the research into IVAs addresses agents with a 'positive' attitude towards the user; i.e., they aim to support or collaborate with her in one way or the other. Instead, the area of IVAs with a 'negative' attitude has received less attention. This is a missed opportunity, because in our everyday reality, encounters with negative individuals are strikingly common, especially within the public domain [13, 32]. Indeed, in the past few years we can observe an increasing interest in the use of IVAs that also display negative, or even *aggressive* behavior. For instance, recent applications include virtual training of aggression de-escalation skills (e.g., for security personnel or clinical care providers) [6], Virtual Reality exposure therapy [29], and anti-bullying education [36]. In line with these developments, the current paper is part of a project that aims to develop a serious game for aggression de-escalation training in the public domain. The main idea is that trainees will learn to apply appropriate communication techniques by placing them in situations where they have to interact with aggressive virtual characters.

Nevertheless, ensuring that such training applications (and more specifically, the virtual agents involved) achieve the desire effect is a nontrivial issue. In general, it is now commonly accepted that, for an IVA to be effective, it needs to be *believable* [4]. This is particularly true when it comes to aggressive virtual agents, who are expected to induce some 'state of anxiety' in the humans they interact with. After all, in case users (e.g., of an aggression de-escalation training environment) do not experience their virtual conversation partners as scary, such an application will be of little use to prepare the person for aggressive confrontations in the real world.

Hence, an important challenge within this area of research is to develop aggressive virtual agents that induce a state of anxiety in their human conversation partners that approximates the state of anxiety they experience during real world confrontations with aggressive individuals. To reach this (long-term) goal, as a starting point it is useful to know how far we are currently away from it. In other words, to what extent is it already possible to induce the desired state of anxiety by using state-of-the-art virtual agent technology? This is the main purpose of the current paper. An experiment has been set up in which 28 participants were asked to listen to a story told by a conversation partner that would suddenly show an outburst of aggression. For half of the participants, the storyteller was a real human (an actress), whereas for the other half it was a virtual agent. Other than that, the difference between both conditions was kept to a minimum. By measuring physiological responses as well as subjective opinions of the participants, the difference in impact between the real and the virtual aggressive agent could be determined.

The remainder of this article is structured as follows. In Section 2, some background information about our project on aggression de-escalation training is provided. After that, Section 3 summarizes the recent literature on believable agents in the context of aggression de-escalation. Section 4 discusses the concept of anxiety as well as possibilities to induce and measure this emotional state. Next, Section 5 introduces the experiment that is central to our study, and Section 6 presents the results. Section 7 and 8 conclude the paper with a discussion and conclusion.

2. AGGRESSION DE-ESCALATION TRAINING

Aggressive behavior against public service workers (e.g. police officers, ambulance personnel, public transport employees) is an ongoing concern in many countries [13, 32]. Most incidents of aggression are of a verbal nature (e.g., insulting, swearing, intimidating), but in about 10% of the cases the conflicts escalate into physical aggression (e.g., threatening, abusing, robbing). To prepare employees to deal with such aggressive confrontations in an effective manner, dedicated training is required.

An important goal of such training is to help trainees develop their *emotional intelligence*: they should be able to recognize the emotional state of the conversation partner, and act upon this. In this respect, an important factor is the distinction between *reactive* and *pro-active* aggression: reactive aggression is characterized as an emotional reaction to a negative event that frustrates a person's desires, whereas pro-active aggression is the instrumental use of aggression to achieve a certain goal [14]. Based on the observed type of aggression, the employee should either apply a more empathic communication style or a more dominant style [1, 6].

To learn to effectively apply such techniques, employees in different domains within the public sector receive dedicated training in communication, both directly on-the-job and in artificial environments. Since on-the-job training is not considered sufficient because of limited possibilities to create the desired learning scenarios, 'offline' training receives much attention. Such training often uses role-play, where the roles are played either by co-students or by professional actors. Although reasonably successful, these types of training have important drawbacks. First, they are very costly, both in terms of money and time. As a result, the frequency by which they are offered is low. And second, there are large differences in the successfulness of role-play-based training: for some students the learning effect is large, whereas for others it is minimal. In conclusion, existing training is expensive, and hard to tailor to individual needs.

As a complementary approach to real world training, the current paper is part of a project that aims to develop a simulation-based serious game for public service workers, by which they can train their aggression de-escalation skills[1]. Users of the system will be placed in front of a 3D Virtual Reality (VR) environment (see Figure 1) that is either projected on a computer screen or on a head-mounted display[2]. During the training, users will be placed in a virtual scenario in a particular domain (e.g., issuing parking tickets, or selling tram tickets), which involves one or more virtual agents that at some point in time start behaving aggressively (e.g., insulting the tram driver because he is late). The user's task is to de-escalate the aggressive behavior of the virtual agents by applying the appropriate communication techniques. Users will be able to communicate with the agents via multiple modalities (e.g., text, speech, facial expression). Meanwhile, they will be monitored by intelligent software that observes and analyzes the behavior and physiological state (e.g., heart rate, skin conductance, brain activity) of the trainee and provides tailored feedback (e.g., as in [18]). Feedback will consist of two categories, namely hints and prompts on the one hand, and run-time modifications in the scenarios on the other hand. An example of the former would be to inform the trainee that (s)he should use a more empathic communication style, whereas an example of the latter would be to decrease the difficulty level in case the trainee makes many mistakes.

Figure 1. Screenshot of the VR environment used for training.

As explained in the introduction, aggression de-escalation is a prototypical example of a domain in which virtual training can only be successful if the virtual environment used is sufficiently realistic. For the most part, this has to do with the fact that the type of training addressed can be categorized as 'training with anxiety'. For such tasks, previous research (e.g., in the domain of shooting behavior) has pointed out that a certain level of

[1] More information on this project 'Simulation-based Training of Resilience in Emergencies and Stressful Situations' can be found at http://stress.few.vu.nl.

[2] The VR Environment has been developed by IC3D Media (http://www.ic3dmedia.com).

'simulated threat' is a necessary criterion to realize an adequate *transfer of training* from the simulated to the real world [22]. For our training environment, this simulated threat can be achieved by enhancing the believability of the aggressive virtual agents that interact with the user[3]. This topic is discussed in more detail in the following section.

3. BELIEVABLE AGGRESSIVE AGENTS

In 1996, Reeves and Nass put forward their Media Equation theory, stating that people have a tendency to treat computers (and other media) as if they were real humans [25]. This is a promising message for the IVA community, as it confirms that it is - in principle - not unnatural for people to interact with virtual agents, which is a good point of departure for applications that make use of such agents, such as simulation-based training. Nevertheless, the extent to which humans perceive virtual agents as realistic may vary, depending on a number of characteristics. Important concepts in this regard are *believability*, *immersion* and *presence*.

Believability is a property of virtual agents that refers to the extent to which they provide the illusion of being alive, hence permit the audience's suspension of disbelief [4]. In [12], believability is defined by three dimensions, namely *aesthetic*, *functional*, and *social* qualities of agents, which can be related, respectively, to the agent's *body* (its physical appearance), *mind* (the mechanisms that drive its behavior), and *personality* (the traits that determine its interaction style). Believability has some similarities to, but is not the same as immersion. Immersion can be considered a property of virtual environments as a whole (not necessarily of virtual agents), and describes the extent to which the environment is 'capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant' [30]. One way to create an immersive environment is, for instance, to make use of a head-mounted display (HMD). Presence, finally, is defined by Slater and Wilbur as a 'psychological sense of being in the virtual environment' [30], although this is just one out of many existing definitions. As such, presence can be considered a property of a human, rather than of a technology, which is indeed the perspective taken in this paper.

To relate these concepts to the current paper, we aim to study how virtual agents that show believable aggressive behavior in a (possibly immersive) virtual environment affect the state of presence of the human user (with an emphasis on her state of anxiety). Believability of the agents is realized in a number of ways: to create believable aesthetics, we make use of highly realistic 3D virtual characters that are provided by a commercial company. To create believable functional and social abilities, the agents' behavior is generated based on state-of-the-art commercial motion capture technology, which can be used to animate the entire facial motion of the virtual characters and synchronize it with recorded human speech (see Section 5). With the help of a professional actress, pre-recorded speech fragments were created in which the agent shows a sudden outburst of aggression during an otherwise neutral story. Immersion is realized (to some extent) by displaying the virtual character on a 19" computer monitor[4], and providing the auditory modality via a circumaural headphone. Our main interest is to investigate whether this setup results in a strong sense of presence (or more precisely: a sense of anxiety, which can be viewed as a more specific psychological state with a negative valence), and whether this experienced anxiety is similar to the anxiety experienced in the same circumstance in the real world.

Previous work in this area suggests that, under certain conditions, virtual environments and IVAs may indeed trigger physiological, experiential and behavioral responses in humans. For instance, Houtkamp and colleagues [19] found visual dynamics and sounds to be important factors influencing a user's affective appraisal of 3D virtual environments, although their studies did not focus on virtual agents. In contrast, Prendinger et al. [23] did focus on the interaction between humans and virtual agents. They demonstrated, among others, that electrodermal activity (EDA) is a useful indicator for arousal during such interactions. Following up on these developments Choi et al. [10] recently found that affective virtual agents strongly influence people's decision making through a combination of affective and cognitive processes. Additionally, they found physiological evidence that electrodermal lability, an individual trait, predicts which of these processes will dominate. Nevertheless, none of these studies explicitly compared the impact of virtual agents with that of real humans. A welcome exception is the work by Villani et al. [34]. This study indicated that, for the case of a simulated job interview, it is possible to induce physiological and experiential responses in humans that approximate (and on some aspects are even stronger than) the responses triggered by a comparable situation in the real world. However, this study differs from ours in the sense that these authors deliberately added some elements to the virtual condition (such as technological mediation and social and cultural cues) that were not present in the real world condition. In contrast, we aim to set up our experiment in such a way that the virtual and the real condition are as similar as possible.

In sum, despite a body of interesting literature in the area of human-agent interaction and physiological measurements, not many studies compare the impact of virtual agents on (physiological, experiential and behavioral) responses in humans in a systematic manner with the impact of real people; let alone that they focus on the specific impact of aggressive behavior on the spectator's state of anxiety. This is exactly the purpose of the current paper. Before we describe the setup of our experiment, we will provide a more precise definition of anxiety, and discuss possibilities to induce and measure it.

4. ANXIETY

The literature on anxiety mentions a number of closely related concepts, which are used in similar contexts. To avoid confusion on what the targeted mental state is in our experiment, in this section a more detailed explanation of the concept is provided.

Firstly, to reach a high state of presence a trainee needs to experience arousal [3]. In the broadest sense, arousal is considered a physiological and psychological state of alertness and readiness for action, often in response to a particular stimulus.

[3] There are numerous ways in which this 'simulated threat' can be enhanced, e.g. using air blast devices or electric surges. Discussing this in detail however goes beyond the scope of this paper.

[4] No HMD was used in the current study, as this will not be used in the initial version of the aggression de-escalation training system either.

In this paper however, the focus lies on a specific type of arousal underlying the fight-or-flight reaction or acute stress response, where the arousal is specifically caused by a threatening stimulus [8]. This negatively valenced arousal is what we will refer to as (state) anxiety.

A large body of research exists on the physiological effects of anxiety [31], as well as on related work using stress [7] or fear [11] as a denomination for a similar state. Since arousal is an activation of the sympathetic nervous system, electrodermal activity and cardiovascular responses are often used as measurements thereof, as covered in the previous section. For electrodermal activity, tonic changes - changes over longer time-intervals, often referred to as skin conductance levels - are assumed to be indicative of a general level of arousal [15]. Even though cardiovascular activity is influenced by both the sympathetic and parasympathetic nervous system and thus provides a less clear picture of arousal [16], measurements such as heart rate (variability) or blood pressure are reliably used for estimations of arousal [9]. Other modalities that are regularly used as a measurement of arousal include the eye (i.e. pupil size [5]), muscles (i.e. activity [35]) and recently also EEG (i.e. alpha asymmetry [7]).

To induce anxiety using virtual stimuli, a wide variety of methods can be used. For instance, the International Affective Picture Set (IAPS) is composed of a large number of pictures that have been rated on their arousal, valance and dominance [21]. By offering the most negatively valenced and highly arousing pictures to human viewers, a state of negative arousal can be induced which arguably is similar to a state of anxiety. However, the passive nature of viewing still images is at a disadvantage in inducing emotional states compared to viewing video clips [17]. Taking it another step further, videogames are used more and more for emotion elicitation [24]. By using virtual reality in combination with a HMD, presence can be increased further, thus providing a great opportunity for emotion elicitation [28].

As the quality of virtual reality keeps improving, the question raised in the introduction (i.e., 'how far we are currently away from inducing a state of anxiety that approximates the experience of real world confrontations?'), becomes increasingly relevant. In the following sections, as a first step towards answering this question, an experiment and its results are described to compare virtually elicited anxiety with that induced by a real human.

5. EXPERIMENT
The main question to be addressed by the experiment is roughly as follows: does a virtual agent elicit as much anxiety as a human being does during an aggressive outburst towards a human conversation partner? This will be tested in a between-subjects experiment, involving one independent variable (agent type: human or virtual) and several dependent variables that are indicative for anxiety, including physiological (heart rate and EDA) and subjective measurements.

5.1 Preparation
As preparation for the experiment, two video fragments of a virtual agent have been created. To this end, the Faceshift motion capture technology (http://www.faceshift.com/) has been used, in combination with a Microsoft Xbox 360 Kinect camera. The Faceshift software can be used to animate the entire facial motion of virtual characters, including facial expression, head orientation and eye gaze, and synchronize it with recorded human speech. To create the required content for the experiment, a professional actress was hired; she was asked to read a story about the history of the personal computer, while her voice and motion was being recorded. The first fragment lasted about 5 minutes; during this fragment, the actress simply read the text in a calm manner, while making eye contact with the camera from time to time, as if she was reading the text to an audience. After that, a second fragment was recorded. During this fragment, the actress read another part of the story (which followed the first fragment), but after 1 minute she would suddenly get extremely angry towards the camera, while shouting and accusing the listener of not paying attention. The exact text used for this was as follows (translated from her original language): '*What the hell! Just look at me when I'm talking to you, that's the least you can do! It's not difficult, just listen to me, or at least pretend you're listening!*'. After recoding both fragments, they were projected on a 3D virtual character (taken from http://www.rocketbox-libraries.com/) that was selected to resemble the real actress (see Figure 2).

Figure 2. Screenshot of the virtual agent.

5.2 Participants
As participants for the experiment, 28 Computer Science students, all aged between 18 and 25, were recruited. The participants were randomly distributed over two groups of 14 participants, to which we will refer as the 'real' and the 'virtual' condition. The real group consisted of 8 male and 6 female students; the virtual group consisted of 7 male and 7 female students.

5.3 Design
The experiment took place in two ordinary office rooms, of which the physical setup is depicted in Figure 3 (left: real condition; right: virtual condition).

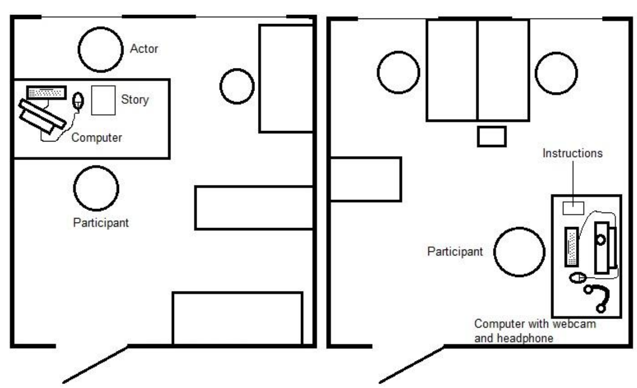

Figure 3. Experimental setup: real (left) and virtual (right).

The rectangles denote tables and the circles denote chairs. In both conditions, the chair that was intended for the participants was placed in front of a table. In the real condition, a real person (the same actress that played the behavior of the virtual agent) was sitting at the other end of the table. In the virtual condition, a computer monitor was placed on the table, on which the virtual agent would be displayed. In this condition, also a (non-working) webcam was attached to the computer, to make it appear as if the participants were monitored by the system (they were told beforehand that the agent was 'intelligent' in the sense that it would use these observations to adapt its own behavior towards the participant).

A summary of the different steps of the experiment is as follows. Before the start of the experiment, the participants had to read a text with instructions and sign an informed consent form. They were told that the experiment focused on measuring their attention and working memory. They were instructed to pay close attention to the story that they were about to hear, since they would have to answer some questions about it afterwards[5]. After that, they entered the experiment room, and took a seat on the chair. The experimenter then attached some physiological sensors (www.plux.info) to their body, which were used to measure their heart rate and electrodermal activity. After testing the sensors, the experimenter left the room.

At this moment, the storyteller (i.e., the human or virtual agent, depending on the condition) started reading the text about the history of the personal computer. As mentioned earlier, the first fragment of this story would take about five minutes. The main goal of this first part was to allow the participant to recover from the initial stress induced by the start of the experiment. After this first part, a short break followed. Then, the second part was read and at this moment physiological measurements were recorded; after one minute, suddenly the aggressive outburst would take place. In the virtual condition, this happened in the way described above. In the real condition, the actress would use the exact same wordings as used for the fragments of the virtual agent. She was also instructed to keep her performance as constant as possible across trials, by always using the exact same intonation, facial and bodily expressions. Then, after the aggressive outburst, the actress would leave the room. In the virtual condition, a 'connection lost' message would appear. This happened approximately 50 seconds after the start of the outburst. A few seconds later (at 1 minute after the outburst), recordings of physiological measurements were stopped and the actress and/or experimenter would re-enter the room, to explain to the participant that the aggressive outburst was just an act and provide comfort when necessary.

After the experiment was terminated, the participants had to fill out a questionnaire. This was done to obtain some information about their subjective experience, in addition to the physiological measurements only. The questionnaire consisted of the following questions:

- Q1: Did you listen carefully when the story was told?
- Q2: Did the aggressive outburst scare you?
- Q3: Although the actual goal of the experiment was to investigate your reaction to aggression, we told you in advance that the experiment focused on attention and memory. Did you believe this?
- Q4: Did you find the aggressive outburst believable?
- Q5: Did you feel personally addressed by the storyteller?

The participants had to rate these questions and score them on a 10-point Likert scale (with 1 = 'absolutely not' and 10 = 'absolutely'). They could also fill out an empty text box for every question, in case they had additional remarks.

6. RESULTS

The results of the physiological measurements are depicted in Figure 4. These graphs depict the dynamics of the heart rate in beats per minute (upper figure) and electrodermal activity in microSiemens (lower figure) over time during the relevant part of the experiment, averaged over all participants in each group. The horizontal axis denotes a period of 2 minutes, i.e., 1 minute before the start of the aggressive outburst and 1 minute after it. The vertical line indicates the moment the outburst started.

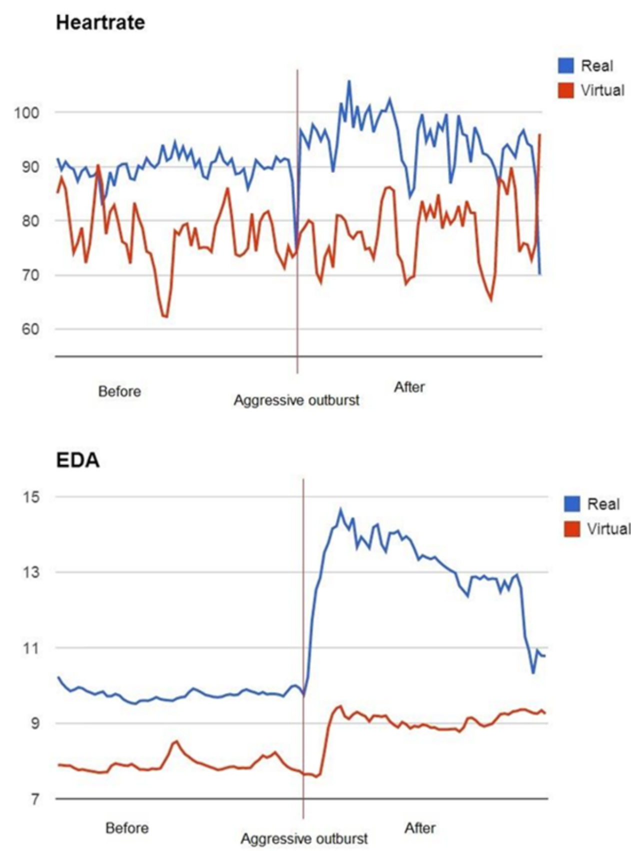

Figure 4. Development of heart rate (above) and EDA (below) during the experiment, averaged over all participants/group.

As can be seen from the figures, it seems that the participants in the real condition already start with a higher level of arousal (on

[5] The participants were asked not to talk back to the (virtual or human) storyteller, as this would result in a too much variation in conversations, which would make the trials incomparable. For the same reason, the actress was instructed to ignore the behavior of the participants.

average) than those in the virtual condition[6]. Indeed, this difference turns out to be significant, both for heart rate and for EDA. During the minute before the outburst, the average heart rate of the real group was 89.7 BPM (σ=10.5), whereas it was 77.3 BPM (σ=11.9) for the virtual group. An independent t-test confirms that this is a significant difference ($p<0.01$). For EDA, the average value of the real group was 10.2 μS (σ=6.0) during that minute, whereas it was 7.9 μS (σ=3.8) for the virtual group. Again, this difference is significant, although the effect is less pronounced than for heart rate ($p<0.05$).

Next, within-group comparisons are made between the average response before and after the aggressive outburst, both for heart rate and the EDA. This is done to find out for each group whether the aggressive outburst caused a significant stress reaction. For the real condition, the average heart rate increased from 89.7 (during the minute before the outburst, as reported above) to 94.8 (during the minute after the outburst). A paired t-test confirmed that this increase is fairly significant ($p<0.05$). The average EDA for this group increased from 10.2 to 13.4, which is strongly significant ($p<0.001$). For the virtual condition, the average heart rate increased only mildly, namely from 77.3 to 78.2; this increase is not significant ($p=0.37$). However, the average EDA for this group did increase significantly ($p<0.001$), namely from 7.9 to 9.0.

The obvious next step is to test whether the increase in responses in the real condition is significantly larger than the increase in the virtual condition. To this end, for both groups and both types of physiological measurements, the differences in response between the minute before and the minute after have been calculated, and these have been compared between the groups using independent t-tests. For heart rate, the average change over time was +5.11 for the real condition, and +0.95 for the virtual condition. Although this difference may seem large, it turned out to be not significant ($p=0.14$), probably due to the large individual differences. For EDA, the average change over time was +3.14 for the real condition, and +1.05 for the virtual condition. In contrast to the situation for the heart rate, this difference is strongly significant ($p<0.001$).

Regarding the subjective measurements, the answers given to all questions in the questionnaire were averaged and visualized in Figure 5. For each question, the difference in ratings between the real and the virtual condition was compared using an independent t-test. There appeared to be a significant difference for question Q1 ($p<0.05$), indicating that the participants in the real condition reported paying significantly more attention to the storyteller than those in the virtual condition. For question Q2, addressing the extent to which the participants experienced being scared, the difference turned out the be just above the significance level ($p=0.13$). For the other three questions, the differences were clearly not significant (all p-values higher than 0.4).

Finally, we tested whether there were any correlations between the answers given in the questionnaire. This was done by calculating Pearson's correlation coefficient for all combinations of questions. For both conditions, there turned out to be moderate to strong relationships between several combinations of questions, most notably between Q2, Q4 and Q5. As an illustration, a scatterplot of the strongest relation, which is between Q2 and Q5 in the real condition ($r=0.61$), is shown in Figure 6. This figure indicates that (for the exception of one outlier) the participants that claimed to feel more personally addressed by the storyteller generally reported to be more scared by the outburst.

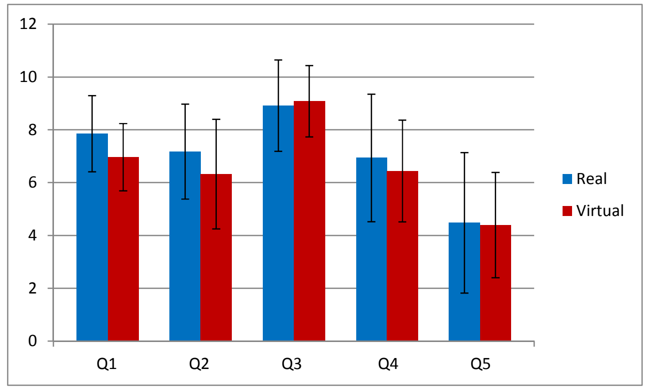

Figure 5. Results of the questionnaire.

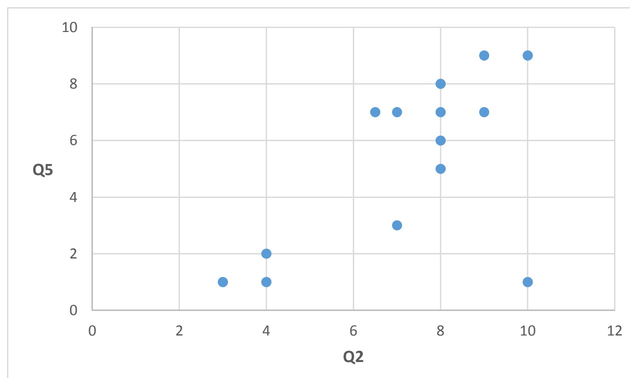

Figure 6. Correlation between Q2 and Q5 (real condition).

7. DISCUSSION

Before discussing the physiological measurements, a discussion on the subjective results is given first. Here, only one result appeared statistically significant, indicating that participants in the real condition reported paying more attention compared to the virtual condition (question Q1). It is not a stretch to consider that there is more pressure on the participant to pay attention when confronted with a real human agent due to social convention. This is consistent with the additional remarks given by some of the participants in the virtual condition, who reported having some difficulties in maintaining a high level of concentration during the experiment. Interestingly however, the other questions did not show any significant results. Thus, no significant differences were found in the scariness (Q2) and the believability of the aggressive outburst (Q4), and in the extent to which the participants felt personally addressed (Q5). One factor that should be taken into account here is a potential bias of the participants in the virtual condition; for instance, these participants may have (unconsciously) interpreted question Q2 as "Did the aggressive outburst scare you, *considering that it was a virtual agent*?".

---

[6] Hence, in a way the two groups have different 'baseline' measurements. To give an impression about the difference with a 'true' baseline measurement: in a pilot study, where subjects were brought in a state at rest by watching a 'relaxing' movie, we found this baseline to have a mean value of 70 BPM for heart rate and of about 7.5 μS for EDA.

Not finding a significant difference in the scariness is interesting, as it might indicate that there is already very little difference between the virtual and the real condition. Also with regard to believability, it is promising that no significant differences were found. However, it is important to note that the sample size in the experiment (28 participants) was limited, which makes it difficult to draw crisp conclusions about this. Additionally, with an average score between 6 and 7 it could be argued that a participant does not experience the same threat in such an experimental condition as it would in a real situation. Nonetheless, this same difference holds for actor-based training versus on-the-job training.

Regarding correlations between the questions, the most pronounced result is the strong correlation between Q2 and Q5, indicating that people who felt more personally addressed by the storyteller generally reported to be more scared by the outburst. This is an interesting finding with our training application in mind. Although correlation does not necessarily mean causation, it may be worthwhile to design the virtual agent in such a way that users have the impression that it addresses them in person.

Next up are the physiological results. A striking first find is that for both the heart rate and electrodermal activity, average values in the real condition are overall higher compared to the virtual group. This might indicate that when confronted with a real human, participants were already more aroused than those who were put in front of a virtual agent. This finding is consistent with the fact that participants in this condition paid more attention to the storyteller (see results for question Q1 in the questionnaire), as attentional effort is typically paired with physiological reactions.

Looking at the impact of the aggressive outburst (by comparing the response levels before and after this moment), a clear effect can be seen in the graph of the electrodermal activity[7]. Even though the aggressive outburst only covers part of the post-measurements, the effect lasts for the entire time in the virtual condition and only at the very end a drop occurs in the real condition. For heart rate the outburst seems to have less impact, although it still resulted in a significant increase in average heart rate in the real condition. This could be caused because cardiovascular measurements not only reflect the sympathetic nervous system, but are of a more complex nature, as explained earlier. A noteworthy result is the short, sudden and extreme drop in heart rate in the real condition at the exact time of the aggressive outburst. This effect is not seen in the virtual group, and could reflect a literal skip of a heartbeat.

To check how sensitive the results are to the time frame used for the physiological measurements, (e.g., due to peculiarities of the specific measures, like EDA having a much slower time decay than heart rate), we investigated whether using measurements over shorter time frames (e.g. 30 seconds before and after the outburst) changed the results, but in general it did not. For instance, using 30 instead of 60 seconds changes the average heart rate before the outburst from 89.7 to 89.8 (see also Fig.4). In summary, physiological measurements seem to indicate that in both conditions, the outburst is sufficient to trigger an emotional response. Both the average level of arousal as well as the reaction to the aggressive outburst are larger in the real condition than in the virtual condition. However, these conclusions should be handled with care, as there was already a difference in the 'baseline' measurement between the two groups. Furthermore, a subjective evaluation of the difference between the two groups with regard to anxiety was inconclusive.

With this in mind, some of the potential weaknesses in this experimental setup seem minor. Even though the virtual agent was designed to be as realistic as possible, some important drawbacks of the fragments need some consideration. First, only the face of the actress was recorded and no gestures were animated in the clips. Second, subtle non-verbal cues displayed in the real world setting, might have gone lost in recording. Third, a regular monitor was used instead of an HMD, which has possibly limited immersion. Nonetheless, the virtual group experienced a physical and emotional reaction to the aggressive outburst. By improving the virtual condition, it can only be expected to decrease differences further, or even surpass the real condition with respect to evoking anxiety.

Although the actress was instructed to keep her performance as constant as possible across participants, it is possible that some small differences appeared, e.g. in the precise tone or volume of her voice or her non-verbal behavior. In any case, similar to the virtual condition, participants in the real condition did not interact with the actress (as instructed) and as a result she was able to consistently give a steady performance, with hardly any interference from the participant. To prevent this discussion entirely, for follow-up studies it may be interesting to add another condition in which participants are confronted with a pre-recorded video fragment of the actress (comparable to a 'teleconference' setting). On the other hand, this would be of less relevance for our project, as in the end the aim is to train people that are confronted with real world, face-to-face aggression.

Another difficult aspect of this type of research is pinpointing or describing the exact emotional state that is being targeted. Here, anxiety was used for the state of negative arousal as described before. In the current setup, there could be a difference between the initial (unconscious) 'shock' experienced when the actress or agent suddenly bursts out in anger compared to the 'higher-order emotion' (of shame perhaps) that is felt afterwards, because of the reprimand. In reality, there is probably a gradual transition from the former state to the latter, which makes it difficult to assess which of the (subjective and physiological) measurements relate to which state. In some sense, this is not a problem for our research, as our main focus is on minimizing the difference between the real and the virtual condition, for all possible responses. At some point however, it might become useful to gain a deeper understanding about the relation between the different measurements and the emotional states that are induced.

Related to this, an interesting point of discussion addresses the question which aspects of the (real and virtual) storyteller's aggressive behaviour contribute to the state of anxiety. Potential components include facial expression, sound, and the content of the utterances. For instance, it is possible that the main part of the participants' reaction (both in the real and the virtual condition) was caused by an acute stress response triggered by the loud sound of the shouting storyteller, independent from what she said and the look on her face. In follow-up experiments, we aim to study these aspects in more detail.

[7] The slight delay of the response in the virtual condition appeared to be caused by a timing issue in the fragments of the virtual agent.

8. CONCLUSION

In this paper, an experiment was described to compare the impact of an aggressive virtual agent with that of a real human actor with respect to the listener's anxiety. A group of 28 participants took part in the experiment, which were randomly assigned to the 'real' or the 'virtual' condition. Results show a physiological response to the aggressive outburst in both conditions, although it is larger in the real group. In contrast, a post-activity questionnaire did not show any differences in subjectively experienced anxiety between both groups. In all, these findings show that virtual agents are able to elicit similar responses as a real human actress, when showing aggressive behavior. Although there are still some differences in the physiological response induced, this provides strong support for using virtual agents for aggression de-escalation training.

With multiple possibilities to enhance the realism of the virtual agent, in the future the emotion experienced by the user could be intensified even further, which provides more possibilities into exploring the differences (or lack thereof) between the real and the virtual world. Currently, a virtual environment for training aggression de-escalation techniques is being implemented, which will be used to investigate the potential of using virtual agents for training compared to role-play further. Here, in addition to the currently used facial animations, agents will be able to make gestures. Moreover, in addition to evaluating the impact on users' emotional and physiological state, transfer of learning from the virtual training environment to the real world will be evaluated.

Parallel to this, it is important to also consider the possible risks of using virtual reality. As it is not yet known what the limits are that can be reached using this technology, care should be taken not to reach a situation in which the virtual experience becomes equally (or more) frightening than real-life, hence introducing similar (psychological) risks such as posttraumatic stress disorder. As this research has shown, there lies great power in using virtual reality and virtual agents for training purposes, but it is important we also learn to understand and harness this power well.

ACKNOWLEDGEMENTS

This research was supported by funding from the National Initiative Brain and Cognition, coordinated by the Netherlands Organization for Scientific Research (NWO), under grant agreement No. 056-25-013. The authors are very grateful to Maaike Schuilenburg for her contribution to the experiment. Additionally, they wish to thank Marco Otte for his technical support, as well as all students who participated in the experiment.

9. REFERENCES

[1] Anderson, L.N. and Clarke, J.T. (1996). De-escalating verbal aggression in primary care settings. Nurse Pract. 21(10):95, 98, 101-2.

[2] Aylett, R., Krenn, B., Pelachaud, C., and Shimodaira, H. (eds.) (2013). Intelligent Virtual Agents. Proceedings of the 13th International Conference on Intelligent Virtual Agents, IVA'13. Lecture Notes on Computer Science, Springer Verlag, vol. 8108.

[3] Baños, R.M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., Rey, B. (2004). Immersion and Emotion: Their Impact on the Sense of Presence. CyberPsychology & Behavior 7(6), 734–741.

[4] Bates, J. (1994). The role of emotions in believable agents. Communications of the ACM, vol. 37, issue 7, pp. 122-125.

[5] Bradley, M., Miccoli, L., Escrig, M., and Lang, P. (2008). The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology , 45, 602-607.

[6] Bosse, T. and Provoost, S. (2014). Towards Aggression De-escalation Training with Virtual Agents: A Computational Model. In: Proceedings of the 16th Int. Conference on Human-Computer Interaction, HCI'14. Lecture Notes in Computer Science, Springer Verlag, pp. 375-387.

[7] Brouwer, A.M., Neerincx, M.A., Kallen, V.L., van der Leer, L., ten Brinke, M. (2011). EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress. Journal of Cybertherapy & Rehabilitation 4(1), 21–34.

[8] Cannon, W. (1932). Wisdom of the Body. United States: W.W. Norton & Company.

[9] Chida, Y. and Hamer, M. (2008). Chronic psychosocial factors and acute physiological responses to laboratory-induced stress in healthy populations: a quantitative review of 30 years of investigations. Psychological Bulletin, 134 (6), 829-885.

[10] Choi, A., Melo, C.M. de, Khooshabeh, P., Woo, W., and Gratch, J. (2015). Physiological evidence for a dual process model of the social effects of emotion in computers. International Journal of Human-Computer Studies 74, pp. 41-53.

[11] Courtney, C. G., Dawson, M. E., Schell, A. M., Iyer, A., and Parsons, T. D. (2010). Better than the real thing: Eliciting fear with moving and static computer-generated stimuli. International Journal of Psychophysiology, 78(2), 107-114.

[12] De Angeli, A., Lynch, P., and Johnson, G. (2001). Personifying the e-market: A framework for social agents. In: M. Hirose (Ed.), Proceedings of Interact 2001. IOS Press, pp. 198-205.

[13] DelBel, J. (2003). De-escalating workplace aggression. Nurning Management (USA) 34, 31-34.

[14] Dodge, K.A. (1990). The structure and function of reactive and proactive aggression. In D. Pepler and H. Rubin, (eds.), The development and treatment of childhood aggression (pp. 201-218). Hillsdale, NJ: Erlbaum.

[15] Figner, B. and Murphy, R. (2011). Using skin conductance in judgment and decision making research. In M. Schulte-Mecklenbeck, A. Kuehberger, & R. Ranyard (Eds.), A handbook of process tracing methods for decision research (pp. 163-184). New York: Psychology Press.

[16] Geus, E. de and Doornen, L. van (1996). Ambulatory assessment of parasympathetic/sympathetic balance by impedance cardiography. In J. Fahrenberg, & M. Myrtek, Ambulatory Assessment. Computer assisted psychological and psychophysiological methods in monitoring and field studies (pp. 141-164). Berlin: Hogrefe & Huber.

[17] Gross, J.J. and Levenson, R.W. (1995). Emotion elicitation using films. Cognition & Emotion, 9(1), 87-108.

[18] Heuvelink, A. and Mioch, T. (2008). FeGA: A cognitive Feedback Generating Agent. In Proceedings of the Seventh

IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2008) (pp. 567-572). IEEE Computer Society Press.

[19] Houtkamp, J.M., Schuurink, E.L., and Toet, A. (2008). Thunderstorms in my computer: the effect of visual dynamics and sound in a 3D environment. In: Bannatyne, M. & Counsell, J. (Eds.), Proc. of the Int. Conference on Visualisation in Built and Rural Environments BuiltViz'08, IEEE Computer Society, Los Alamitos, USA.

[20] Kim, J., Hill, R.W., Durlach, P., Lane, H.C., Forbell, E., Core, C., Marsella, S. Pynadeth, D. and Hart, J. (2009) BiLAT: A game-based environment for practicing negotiation in a cultural context. International Journal of AI in Education, vol. 19, issue 3, pp. 289-308.

[21] Lang, P.J, Bradley, M.M., and Cuthberth, B.N. (1999). International Affective Picture System (IAPS): Technical Manual and Affective ratings. Gainesville, Fl. The Cennter for Research in Psychophysiology, University of Florida

[22] Nieuwenhuys, A. and Oudejans, R.R.D. (2011). Training with anxiety: Short- and long-term effects on police officers' shooting behaviour under pressure. Cognitive Processing, 23 March 2011, 1-12.

[23] Prendinger, H., Becker-Asano, C., and Ishizuka, M. (2006). A study in users' physiological response to an empathic interface agent. International Journal of Humanoid Robotics, 3(3):371–391.

[24] Reekum, C. van, Johnstone, T., Banse, R., Etter, A., Wehrle, T., and Scherer, K. (2004). Psychophysiological responses to appraisal dimensions in a computer game. Cognition and Emotion, 18(5), 663-688.

[25] Reeves, B. and Nass, C. (1996). The media equation: How people treat computers, television, and new media like real people and places. Stanford, CA : CSLI Publications.

[26] Rickel, J. (2001). Intelligent virtual agents for education and training: Opportunities and challenges. In A. de Antonio, R. Aylett, and D. Ballin (eds.), Intelligent Virtual Agents, Lecture Notes in Computer Science, Springer Verlag, vol. 2190, pp. 15–22.

[27] Ritterfeld, U., Cody, M., and Vorderer, P. (eds.) (2009). Serious Games: Mechanisms and Effects. New York/London: Routledge.

[28] Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, Gaggioli, A., D., Botella, C. and Alcañiz, M. (2007). Affective interactions using virtual reality: The link between presence and emotions. CyberPsychology & Behavior, 10(1), 45-56.

[29] Rizzo, A.A., Reger, G., Gahm, G., Difede, J., and Rothbaum, B.O. (2008). Virtual Reality Exposure Therapy for Combat Related PTSD. In: Shiromani, P., Keane, T., and LeDoux, J. (eds.), Post-Traumatic Stress Disorder: Basic Science and Clinical Practice, Springer Verlag.

[30] Slater, M. and Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. Presence: Teleoperators and Virtual Environments, 6(6).

[31] Smith, J.C., Bradley, M.M., and Lang, P.J. (2005). State anxiety and affective physiology: effects of sustained exposure to affective pictures. Biological Psychology, 69(3), 247-260.

[32] Smith, M.J. and Clarke, R.V. (2000). Crime and public transport. In R. Clarke (Ed.), Crime and Justice: A Review of Research (pp. 169-234). Chicago, IL: The University of Chicago Press.

[33] Traum, D. and Rickel, J. (2002). Embodied agents for multi-party dialogue in immersive virtual worlds. Proceedings of the International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2002), pp. 766–773.

[34] Villani, D., Repetto, C., Cipresso, P., and Riva, G. (2012). May I experience more presence in doing the same thing in virtual reality than in reality? An answer from a simulated job interview. Interacting with Computers, 24 (4), 265-272.

[35] Visser, B., Looze, M. de, Graaff, M. de, and Dieën, J. van (2004). Effects of precision demands and mental pressure on muscle activation and hand forces in computer mouse tasks. Ergonomics , 47 (2), 202-217.

[36] Zoll, C., Enz, S., Schaub, H., Aylett, R., and Paiva, A. (2006). Fighting Bullying with the Help of Autonomous Agents in a Virtual School Environment. In: Proceedings of the 7th International Conference on Cognitive Modelling (ICCM).