



Virtually Bad:

A Study on Virtual Agents that Physically Threaten Human Beings

Applicants

Eligible proposals must have two (and only two) applicants from different disciplines within the Network Institute.

Supervisor Name	Department/Group	Faculty
1. Tibor Bosse	Computer Science	FEW
2. Tilo Hartmann	Communication Science	FSW

Project description

Provide a brief description of the project (max. 300 words)

Intelligent Virtual Agents (IVAs), interactive human-like characters¹, become widely used for numerous applications, varying from healthcare decision support² to communication training³. In such applications, IVAs play various roles in which they interact with users, for instance as an instructor or teammate⁴. Interestingly, in the vast majority of these cases, IVAs are friendly and supportive. Instead, the area of IVAs with a 'negative' attitude towards users (i.e., 'virtual bad guys') has been heavily under-researched.

However, 'virtual bad guys' are a highly interesting topic of study for at least two reasons:

- 1) Several prominent people recently expressed their concern that autonomous systems might evolve to a point where they threaten human beings. Controlled studies can provide a better understanding of how humans would react to such threatening AI systems.
- 2) The concept of virtual bad guys opens up a range of useful applications, including virtual training of aggression de-escalation skills (e.g., for security personnel), Virtual Reality exposure therapy, and anti-bullying education.

¹ Brinkman, W.P., Broekens, J., and Heylen, D. (2015). Intelligent Virtual Agents. 15th International Conference, IVA 2015, Delft, The Netherlands, August 26-28, 2015, Proceedings. Springer LNAI.

² DeVault, D., et al. (2014). SimSensei kiosk: a virtual human interviewer for healthcare decision support, Proceedings of the 13th international conference on Autonomous agents and multi-agent systems, AAMAS'14, pp. 1061-1068.

³ Ben Youssef, A., Chollet, M., Jones, H., Sabouret, N., Pelachaud, C., and Ochs, M. (2015). Towards a Socially Adaptive Virtual Agent, Intelligent Virtual Agents, Vol. 9238, pp. 3-16.

⁴ 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.

However, *believability*, a common problem in the design of IVAs, poses a particular challenge for ‘virtual bad guys’. Many effective applications require that users feel indeed seriously threatened or stressed by the IVA. However, as a general problem, IVAs are non-consequential, i.e., are unable to apply serious sanctions to users. Accordingly, users also perceive and categorize IVAs as virtual beings. These factors plausibly shape –and skew– how users respond to IVAs. So how to build virtual bad guys that are taken seriously? The present project tackles this question by designing and experimentally examining the effects of a technologically advanced IVA that is able to seriously threaten users. The study will test effects on stress and anxiety, and assess the role of believability. Ethical boundaries and implications will represent an important part of this project, too.

Project Organization

Each proposal requests two Academy Assistants from different disciplines. Describe their roles and describe the skills and expertise required from them. (max. 300 words)

The main task of the first Academy Assistant (AA) is the **development of the IVA**. This will mainly involve programming in a platform like Unity 3D. Users should interact with the IVA using free speech. The IVA will autonomously turn-take, and will respond by playing pre-defined animations and speech. This content will be recorded in advance, in collaboration with the second AA.

To be able to implement the IVA, the first AA needs to have a background in Computer Science or Artificial Intelligence, and requires good programming skills. Also some affinity with human-computer interaction is desirable.

Additionally, the implementation part could benefit from the support by Marco Otte, who has previously collaborated with both of the applicants in the context of several research projects involving IVAs.

The main task of the second AA is understanding the **media-psychological implications** of the designed IVA. An experiment will compare the effects of the IVA in a standard vs. consequential condition on users’ (physiological and self-reported) stress and anxiety levels. In the consequential condition, users are equipped with a (non-functioning) device, and made believe that through this device the IVA, if intending, is able to apply mildly painful electroshocks. The study will also examine the mediating role of believability and Presence experiences on induced stress levels.

To be able to accomplish these tasks, the second AA needs to have a background in Communication Science or Media Psychology, and requires good analytical and empirical skills. Interest in virtual realities is desirable.

Finally, an investigation of the **ethical implications** and boundaries of the IVA and experiment represents a substantial element of the project. This task includes approval of the experiment by an ethical committee.

Collaboration

Describe how your research improves collaboration and cross-pollination between the disciplines involved (max. 300 words)

Studying the impact of ‘virtual bad guys’ on users’ stress response requires expertise in computational as well as communication science. The project will build on the expertise available at the Computer Science department on the development of IVAs, in particular in the context of human-agent conversations in threatening circumstances (e.g., for aggression de-escalation training⁵).

Communication Science will contribute expertise on understanding/evaluating the (media)psychological impact of such ‘conversational’ agents on humans. It also provides knowledge about believability perceptions and stress induction. The project closely links to work in the Communication Science department on users’ perception of social entities in TV, video games, and VR environments⁶.

Hence, the project serves to more closely integrate computer science and media-psychology thinking in a relevant context. Neither side can be successful without the other:

Computer Science will benefit from this collaboration because it provides opportunities to systematically evaluate the effect of the developed technology on human beings. Such an extensive evaluation goes far beyond the type of testing typically done in Computer Science, which mostly is at the level of simple ‘usability studies’ involving only a handful of participants.

Communication Science benefits, because examinations of IVAs in this discipline often suffer from poor applications and lacking ‘technological’ expertise. Collaborating on designing and testing an IVA provides the unique opportunity to not only examine a well-tailored and suitable IVA, but also to further understanding of the ‘mechanical basis’ and complexity underlying virtual agents – also in the specific context of ‘virtual bad guys.’

Deliverables

Enumerate intended project results: papers, research proposals or otherwise. (max 200 words)

The project is expected to lead to a number of results:

- A prototype of a ‘virtual bad guy’, i.e. an interactive system that enables a user to have a natural conversation with an IVA, which has the ability to physically threaten the user and is perceived as ‘sufficiently serious’.
- The design of an experiment that can be used to test the psychological impact of the system on human users.
- A report in which the results of the experiment are documented.
- A joint paper presented at an international conference (incl. proceedings) and/or submitted to a journal in the area of human-computer interaction or a related discipline.
- If the results are promising, they may be used to write a proposal to acquire funding for follow-up research (e.g., within a call on Creative Industry or Applied Gaming).

⁵ See <http://stress.few.vu.nl/>.

⁶ Hartmann, T., & Goldhoorn, C. (2011). Horton and Wohl revisited: Exploring viewers’ experience of parasocial interaction. *Journal of Communication*, 61, pp. 1104-1121.

Planning

Provide a breakdown of the project into phases with tentative timing (max 150 words)

Phase	Month	AA1 (Computer science)	AA2 (Communication Science)
1	1	Making conceptual design for IVA (i.e., defining scenarios and interaction possibilities)	Exploring relevant literature on IVA believability, Presence, and stress induction in interpersonal context
2	2	Scrutinizing ethical implications	Scrutinizing ethical implications
3	3-5	Implementing the IVA	Designing the experiment; recruiting participants
4	6	Conducting the experiment	Conducting the experiment
5	7-8	Analyzing the results	Analyzing the results
6	9-10	Writing joint paper	Writing joint paper

Please respect the word count limits: proposals that exceed the stated limits will not be eligible.

Send completed proposals to: akademiestudent.fsw@vu.nl, before 10 June 2016 at 12.00pm. An independent committee will evaluate the proposals; subsequent notification of the committee decision will be given on 1 July.