

When a Virtual Agent is a Flawed Stimulus

Extended Abstract

Hannes Högni Vilhjálmsson

Center for Analysis and Design of Intelligent Agents (CADIA)

Reykjavik University, Iceland

hannes@ru.is

One fascinating aspect of IVAs is the potential to bring the benefits of the "human touch" to situations where using real humans is impractical or impossible for some reason. For example, an intelligent tutoring system running on a personal laptop could provide a virtual teacher, an online web store could provide a virtual sales agent and a mobile personal training app could provide a virtual coach. For certain situations, there are undeniable benefits in having a real human at hand, and therefore it may not be such a stretch to propose that virtual humans could provide some of those same benefits. Especially since these virtual counterparts can look quite appealing and capable of generating what seems very life-like behavior.

However, the HCI community does not simply embrace the idea of virtual interface agents without evidence of the benefits over traditional interfaces. After all, having a chat with an animated agent seems in direct opposition to the pervasive guiding principle of direct manipulation in interface design [3, 4]. Therefore, a great number of studies have been conducted where animated agents have been added to various interfaces. While users of these agents were expected to reap the benefits of getting a face-to-face treatment, results have generally been underwhelming, in particular with regards to performance outcomes [1, 2]. From reviewing the literature, it is easy to come to the conclusion that virtual agents are great fun, but add little "real" functionality to HCI.

One of the hardest things about demonstrating how a virtual agent is capable of serving a human-like purpose in a particular scenario, such as in tutoring, is that this is a two-fold challenge:

- (1) The virtual agent itself needs to be a sufficiently good realization of the theoretical behavioral model.
- (2) The virtual agent needs to be able to serve its human-like purpose in the given scenario.

The real difficulty for the field lies in the existence of (1), that is, if the artifact we create is flawed in some way that compromises the human principle we are trying to encapsulate, what comes out of part (2) is completely useless for claiming anything of substance.

For example, let's say the goal is to study whether a particular piece of information is more effectively communicated by a virtual agent than by reading that same piece of information from the pages of a book. For the virtual agent condition, it would not be enough to feed the text to a TTS hooked up to a lip-synced face. The issue is that the agent is meant to represent *embodied communication*, but by being strictly text and lip focused, the realization ignores the important role of nonverbal behavior in that process, both its crucial propositional and interactional components. In essence, the stimulus in the study does not sufficiently capture a basic model of embodiment.

This flaw will be fairly obvious those that study Embodied Conversational Agents, but this is an extreme example for the purpose

of demonstration¹. Usually the virtual agents are more complex and capable of well coordinated multi-modal behavior. But how can we be sure that our realization is sufficiently good, before we put it to the test in our usage scenario? This may require us to address questions like:

- Are all crucial parts of the theoretical model represented in the computational model?
- Are the behaviors generated by the computational model correctly recognized by a human?
- Are there any unintended behaviors, possibly recognized and interpreted by a human?
- Are there any breaks in behavior or missing behaviors, possibly recognized and interpreted by a human?
- Are there any general usability issues?

Many of these questions can be addressed with properly conducted validation tests (e.g. for testing the correct recognition of behaviors) and pilot studies (e.g. where usability issues can often be caught and fixed).

But some questions pose quite a challenge, for example there are no established methods for testing for the correct translation of a theoretical model of social behavior into a computational model. Also, what should be done about those aspects of the agent's behavior that are not governed by the model? If those are completely left out, the absence may trigger an unwanted reaction or interpretation by the human. This could be addressed by providing at least some baseline neutral behavior that the agents always reverts to when not controlled by the model, but what would constitute neutral behavior?

Finally, all of this has serious implications for the replicability of these studies. The focus of many of the reported results is on part (2) of the two-fold challenge, sometimes leaving out crucial bits about the stimulus itself, the virtual agent.

REFERENCES

- [1] Doris M. Dehn and Susanne Van Mulken. 2000. The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies* 52, 1 (2000), 1 – 22. <https://doi.org/10.1006/ijhc.1999.0325>
- [2] Steffi Heidig and Geraldine Clarebout. 2011. Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review* 6, 1 (2011), 27–54.
- [3] Ben Shneiderman. 1983. Direct manipulation: A step beyond programming languages. *Computer* 8 (1983), 57–69.
- [4] Ben Shneiderman. 2010. *Designing the user interface: strategies for effective human-computer interaction*. Pearson Education India.

¹Although this kind of a study has been conducted for real!