Ethical Considerations in Designing Adaptive Persuasive Games

Christoffer Holmgård Pedersen, Rilla Khaled, and Georgios N. Yannakakis

IT University of Copenhagen, Computer Games Research Group, Rued Langgaards Vej 7, 2300 København S, Denmark
holmgard@itu.dk

Abstract. In this poster, we describe an ongoing project concerning the development of an Adaptive Treatment Game (ATG) for treating Post Traumatic Stress Disorder. The ATG uses biofeedback and computer game technology to enable multiple treatment techniques and goals. We examine how a multidisciplinary approach shaped the prototype and we discuss the ethical implications of creating a self-adaptive, semi-autonomous treatment game.

Introduction. Post Traumatic Stress Disorder (PTSD) can be a severely disabling syndrome. It is sometimes developed after exposure to extreme stress in situations that include experiencing or witnessing mortal danger or extreme terror. Research into the efficacy of different treatments for PTSD has been ongoing since the 1980's and a variety of treatment approaches have been identified [2, 8]. One of the most recent developments in treatment approaches is the use of Virtual Reality Therapy (VR-T). Studies of the efficacy of VR-T are cautiously positive, though more research is needed [9].

Meanwhile, advances in affective computing have enabled the creation of systems that use psychophysiological and behavioral data to reliably infer emotions experienced by users, including stress and anxiety [5, 10, 11]. Drawing together threads of earlier research initiatives, we have reason to believe that including ludic and diegetic aspects in VR-T universes will enhance their efficacy, along with their ability to promote attitude and behavior change. To explore this hypothesis, we are developing a prototype of a multi mode Adaptive Treatment Game (ATG) that brings together three Cognitive Behavioral Treatment techniques in one coherent game universe. The ATG prototype will be completed and undergo clinical testing in Spring 2012.

The ATG prototype. The multidisciplinary team behind the ATG included multiple game designers and developers, computer game, affective computing and artificial intelligence researchers and three PTSD therapists (two psychologists and a psychiatrist) with decades of treatment experience between them. Based on the recommendations and experience of the therapists, Relaxation Training (RT), Stress Inoculation Training (SIT) and Exposure Therapy (ET) were chosen as the treatment approaches at the outset of the project. As such,

our tool is multi modal, in that it supports these three treatment types. Avenues of adaptive persuasive design that were outlined by Fogg almost a decade ago [3] have now been used in a plethora of tools and products as discussed by Kaptein et al. [7] and Kaptein and Eckles [6]. Drawing on persuasive design strategies, including *tunnelling*, *tailoring*, and *conditioning* [3], we designed a treatment tool that uses adaptive biofeedback technology to learn an individual patient's response patterns and adjust the presented stimuli relative to reaction data from previous treatment sessions [11]. In addition, the tool uses game design to create a convincing, seamless world. The three modes of the ATG are displayed in Figure 1.

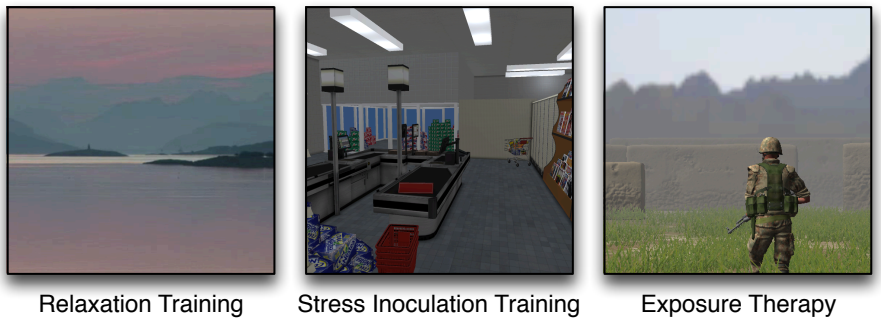

Fig. 1. Screenshots from the three modes of ATG

We decided to create our own development method in order to support the multidisciplinary collaboration process and structure the contributions from the different areas of expertise. Since we wanted to create a game that could be used in real world psychological practice, we needed to ensure that the ATG was feasible, useful and safe outside the laboratory. To solve this task, we started by forming a hierarchy of design concerns, in the following priority: *functional design, treatment design, technology design, and game design*. This *design hierarchy* was used to resolve any design conflicts - e.g. treatment design concerns would always take precedence over game design concerns.

Discussion. A design incorporating input from many sources of reference must become an amalgam of priorities from all the different fields, which are not necessarily compatible. This means that hard decisions and prioritization was necessary in order to make the different constituents of the ATG fit together.

It resulted in an underdeveloped game design, since this was at the lowest tier of the design hierarchy. It might have been fruitful to give game design a higher priority, or to abandon the idea of prioritized concerns altogether to ultimately make a more compelling tool.

However, we believe that the most interesting and pressing questions that the ATG raises, fall under the area of ethical persuasive design. Making any

form of semi-autonomous system that interacts with patients in clinical settings entails a major ethical responsibility on the part of the designers of the system, as does the construction of any piece of persuasive technology. The responsibility of imbuing the system with these adaptive properties is not whisked away by providing the therapist as a safety measure; the constructors of the system still carry a responsibility for its subsequent effects on end users [4]. Berdichevsky and Neuenschwander [1] describe in their decision tree for ethical evaluation of persuasive technologies that a system designer's work is ethical if her system's outcome is intended and good, but she is not responsible if an undesirable outcome is unintended and not reasonably predicable. In the case of adaptive persuasive technology it becomes more difficult to imagine all possible use scenarios and thus all the possible unintended side-effects. This blurs the line of reasonable predictability as also Kaptein and Eckles point out in their treatment of persuasive profiles [6]. Indeed, using adaptivity and profiling might put an even greater responsibility on the designer. In our case, we identified the following risks:

Black-boxing of the ATG's inner workings could make the links between experience and evaluation opaque to the patient and the therapist. This may in term result in alienation from the platform and demotivate the patient from engaging with the ATG more than once. The answer to this was exposing the evaluations of the system to the therapist as well as the patient, making the ATG a tool that the two use in an egalitarian and transparent manner.

Objectification of the patient to a level where the ATG's evaluations take precedence over phenomenological experience. A special responsibility lies with the therapist to emphasize the experience of the patient as valid.

Erroneous profiling where short-comings of the applied AI lead to misclassifications and possible misinterpretations of the patient's reactions to certain stimuli, potentially leading to the exposure of the patient to unduly stressful or completely inappropriate stimuli. This is handled by the fact that the therapist may always override the system.

Second-order conditioning where fear reactions to cues in the virtual environment are not extinguished, but rather generalized, making hitherto unproblematic elements of experience into cues eliciting stress and/or anxiety. This risk is handled in conjunction by the therapist and the ATG.

Re-traumatization could be considered the worst-case consequence of the combination of erroneous profiling and second-order conditioning. If the ATG presented a patient with a wrongly graded, too intense, stimulus, it could set off a fully fledged anxiety attack or a flashback. The consequence could be conditioning adverse responses to the therapy situation itself and have destructive consequences for the therapeutic alliance. To minimize this risk, the stimuli in the ATG undergo testing with expert therapists, users drawn from the general public and as well as veteran cohorts, and carefully selected PTSD patients.

Conclusion and Future Work. With the ATG, we designed and built a prototype that points to a new way of applying virtual reality for PTSD in

particular, but perhaps also cognitive behavioral therapy in general. While we have yet to investigate the efficacy of the ATG as a treatment tool (it will undergo clinical trials in Spring 2012) the process of making the prototype yielded a number of valuable insights.

Bringing a hierarchical set of concerns into an iterative design process turned out to be limiting. With this approach some areas of a project may receive too little attention or be inappropriately bounded by concerns with higher priority. This was partially the case with game design in our project and it remains an open question whether the ATG would be a better tool if game design had been allowed to influence functional or treatment design.

Our research and development efforts so far suggest that adaptive and goal-directed VR-T tools can make psychological therapy not only more engaging, but also more effective at treating debilitating anxiety disorders. It shows that making adaptive and profiling tools raises important ethical questions with responsibilities for the designers and creators – and that handling these challenges is worth the effort, when it allows us to make future cognitive behavioral therapy a more personal, immersive and effective experience.

References

1. Berdichevsky, D., Neuenschwander, E.: Toward an ethics of persuasive technology. Communications of the ACM 42(5), 51–58 (1999)
2. Foa, E.B., Keane, T.M., Friedman, M.J., Cohen, J.A.: Effective treatments for PTSD, practice guidelines from the International Society for Traumatic Stress Studies. New York: Guilford Press, second edn. (2009)
3. Fogg, B.: Persuasive Technology Persuasive Technology. Using Computers to Change What We Think and Do. San Francisco: Morgan Kaufmann Publishers (2003)
4. Friedman, B., Kahn Jr, P.: Human values, ethics, and design. In: The human-computer interaction handbook. pp. 1177–1201. L. Erlbaum Associates Inc. (2002)
5. Haag, A., Goronzy, S., Schaich, P., Williams, J.: Emotion recognition using bio-sensors: First steps towards an automatic system. Affective Dialogue Systems pp. 36–48 (2004)
6. Kaptein, M., Eckles, D.: Selecting effective means to any end: Futures and ethics of persuasion profiling. Persuasive technology pp. 82–93 (2010)
7. Kaptein, M., Markopoulos, P., de Ruyter, B., Aarts, E.: Can you be persuaded? individual differences in susceptibility to persuasion. Human-Computer Interaction–INTERACT 2009 pp. 115–118 (2009)
8. Nemeroff, C., Bremner, J., Foa, E., Mayberg, H., North, C., Stein, M.: Posttraumatic stress disorder: a state-of-the-science review. Journal of Psychiatric Research 40(1), 1–21 (2006)
9. Parsons, T., Rizzo, A.: Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. Journal of Behavior Therapy and Experimental Psychiatry 39(3), 250–261 (2008)
10. Picard, R.: Affective Computing. MIT Press, second edn. (2000)
11. Popović, S., Horvat, M., Kukolja, D., Dropuljić, B., Ćosić, K.: Stress inoculation training supported by physiology-driven adaptive virtual reality stimulation. electronic form only:: NE (2009)