



VU Research Portal

A serious game for training verbal resilience to doorstep scams

van der Lubbe, Laura M.; Gerritsen, Charlotte; Formolo, Daniel; Otte, Marco; Bosse, Tibor

published in

Games and Learning Alliance
2019

DOI (link to publisher)

[10.1007/978-3-030-11548-7_11](https://doi.org/10.1007/978-3-030-11548-7_11)

document version

Publisher's PDF, also known as Version of record

document license

Article 25fa Dutch Copyright Act

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van der Lubbe, L. M., Gerritsen, C., Formolo, D., Otte, M., & Bosse, T. (2019). A serious game for training verbal resilience to doorstep scams. In H. Söbke, M. Gentile, & M. Allegra (Eds.), *Games and Learning Alliance: 7th International Conference, GALA 2018, Palermo, Italy, December 5–7, 2018, Proceedings* (pp. 110-120). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 11385 LNCS). Springer - Verlag. https://doi.org/10.1007/978-3-030-11548-7_11

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl



A Serious Game for Training Verbal Resilience to Doorstep Scams

Laura M. van der Lubbe¹(✉), Charlotte Gerritsen², Daniel Formolo¹,
Marco Otte¹, and Tibor Bosse³

¹ VU Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
{l.m.vander.lubbe,d.formolo,m.otte}@vu.nl

² NSCR, De Boelelaan 1077, 1081 HV Amsterdam, The Netherlands
cgerritsen@nscr.nl

³ Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands
t.bosse@ru.nl

Abstract. People are frequently confronted with scams; swindlers trying to gain your trust to get hold of your personal information, money or belongings. Elderly people are especially vulnerable to these tricks, that typically occur at the front door, street or on the phone. We have developed a virtual training environment to teach people how to handle these situations and learn them to increase their verbal resilience. The training is implemented as a tablet application and consists of six scenarios that are likely to occur in daily life. Participants are placed in a dialogue with a virtual character and may interact by choosing an answer from a fixed multiple-choice menu or by speaking the answer aloud. A speech recognition module is able to detect the level of assertiveness and provides immediate feedback to the user's performance. In this paper, we present the implementation of the virtual training application. To evaluate the prototype a focus group was organized, consisting of potential end users. The outcomes were mainly positive and the provided feedback will be incorporated into the final version.

Keywords: Serious gaming · Learning · User study

1 Introduction

When a con artist tells you a convincing but fraudulent story in order to enter your house and/or rob you this is called a doorstep scam. Doorstep scams frequently happen, numerous news reports about different stories exist. Since doorstep scams often have a high (emotional) impact, various campaigns try to educate people on this topic in order to prevent doorstep scams from happening (e.g. 'Spot it, Stop it'¹). Such campaigns focus at behavioral aspects of the prevention of doorstep scams. Doorstep scams are acknowledged by the Dutch ministry of Safety and Justice as high impact crimes. Because of this,

¹ www.ageuk.org.uk/information-advice/money-legal/scams-fraud/doorstep-scams/.

and because existing campaigns are not enough, they funded research towards a virtual doorstep scam resilience training. This research is executed in collaboration with a large Dutch elderly organization called KBO-PCOB². In a previous paper [8], we proposed the idea of a serious game for training of verbal resilience to high-risk doorstep scams. That paper describes how a virtual training can be used to improve people in their verbal resilience in order to prevent doorstep scams from having negative outcomes. The target group for this virtual training is high-risk victims; for doorstep scams these are elderly people. A training with virtual agents offers multiple advantages, among which are the low costs and the repeatability, comparing it to training with actors. This has already been shown in research in other domains, such as [1, 2].

The current paper describes the implementation of the serious game and is organized as follows: First, an overview of related work is given, followed by a detailed description of the different aspects of the implementation. A first evaluation of the prototype is described in Sect. 4. The paper ends with the conclusion and discussion.

2 Related Work

Doorstep scams mostly happen at the front door, on the street, or on the phone. The goal of a doorstep scam is often to rob or get sensitive information from the victim. The content of doorstep scams is studied in [8]. A list of frequently encountered doorstep scam stories was created, based on a field research, which lies at the basis of the created scenarios. From all the discussed scenarios, the six most encountered scenarios are chosen for the application. Two scenarios take place at the front door, two scenarios take place on the street, and two scenarios take place on the phone. The scenarios are described in more detail in [8].

A specific skill that is beneficial for the resilience against doorstep scams is assertiveness: behaving confident and daring to say what you think or believe³. Various interventions exist to improve people in their assertiveness. Increasing the refusal of an unreasonable request can be done by verbal modeling and therapist coaching [9]. Being able to refuse such a request is an important skill for doorstep scams. Refusing can be done in various ways, among which are simply saying no, but also refusing by changing the subject [12]. Both the different verbal strategies that can be used as well as the need to be assertive while refusing, need to be present in a prevention program in order to teach students to resist indirect and direct pressures to engage in negative behaviors [10]. It is important that the verbal strategies are practiced in specific situations, although nonverbal assertive skills can be used for different situations [12].

Being assertive is also reflected in the way someone speaks, which will also be addressed in the application that is built. For example, speaking firmly or authoritative [11] or the medium latency of the response [4] are associated with an assertive way to use your voice.

² www.kbo-pcob.nl.

³ Definition: dictionary.cambridge.org/dictionary/english/assertive.

However, so far no comparable interventions for elderly people have been found to improve their resilience against (doorstep) scams. Serious games for elderly are often exergames, games in which the player has to perform some sort of physical activity. These games promote active aging or aim to help people with physical problems, such as balance or postural control problems with the use of Nintendo Wii Fit, together with the Balance Board (e.g. [7]) or the Xbox Kinect (such as [14]). Besides exergames, there are serious games for elderly people that help to improve their cognitive abilities, via brain training games such as Smart Thinker [3]. Other serious games for elderly aim to promote social activities, e.g. SilverGame [13].

3 Implementation

The three dimensional environment, which is described in more detail in Sect. 3.2, as well as the user interface for the application are created in the game-engine Unity⁴. In order to program the application various C# scripts are written.

The application starts with a main menu with four different options: scenarios, scores, explanation, and credits. On the scores pages the player can find their top ten scores and the average number of stars received for each scenario. It is also possible to reset all the scores. How the scores are calculated is described in Sect. 3.3. The scenarios, both their content as their functioning, are described in Sect. 3.1.

3.1 Scenarios

The player can choose from six scenarios described in [8], these are played from a first person perspective. The gender of the avatar is chosen randomly by the application. Besides this, the player can choose to play using speech analysis or not, in order to train the assertiveness of his/her voice. The speech analysis is only available when playing with a network connection and is, if connected, by default turned on. The speech analysis is further explained in Sect. 3.4.

Figure 1 shows the flow of the application when playing a scenario. First, the scenario is set up: the camera is moved to the correct location in the environment, and the gender of the avatar and/or voice are randomly chosen. When the scenario needs an avatar (this is the case when it is not a phone scenario), this avatar is activated. Finally, the intro animation is played. For each type of scenario (door, street or phone) another intro is used. In the case of the door scenario the doorbell is rang, the door is opened and the camera moves a bit forward. For the street scenarios the camera and an animated dog, placed close to the camera to represent the players dog, move towards the avatar. In case of the phone scenarios a ring tone is played after which the screen of the phone placed in the environment changes, representing an incoming call.

The scenario always starts with the virtual opponent, the speech is played together with face and body animations (if needed). After the turn of the virtual

⁴ www.unity3d.com/.

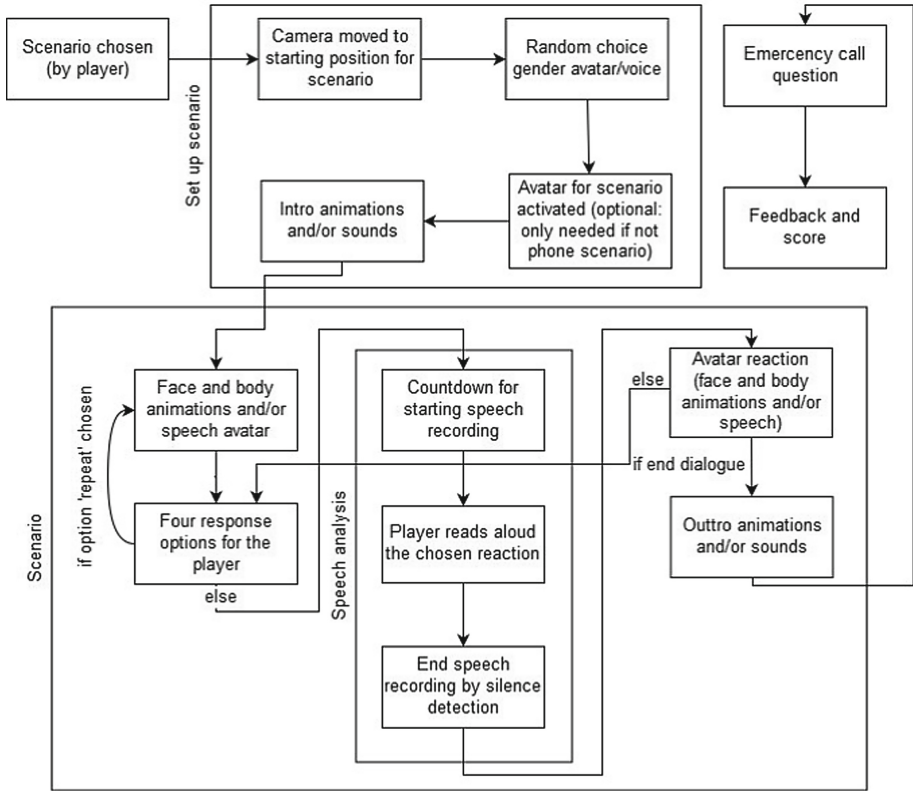


Fig. 1. Flowchart of the application when playing a scenario

opponent the player generally has four possible responses to choose from. One of these responses is to repeat the last turn (this response is only available if the virtual opponent has said something in the previous turn). This response is added to accommodate the target user group. The other responses influence the progress of the scenario.

Figure 2 shows a screenshot of the application when the player has reached a choice moment, during the energy scenario at the door. At the bottom of the screen the four possible reactions are visible, the bottom one is the repeat option, the other three options are randomly ordered. In the upper right corner the gauge for the speech analysis, showing the players last speech analysis score.

The scenarios can be represented in flowcharts, Fig. 3 shows such a flowchart for a scenario. The blue rectangles show the avatar’s dialogue, the red rectangle shows a negative end state with the dialogue of the avatar. The rounded rectangles show the possible reactions of the player, the color of the rectangle represents the rating of the response. For some reactions of the player there are two outgoing arrows with conditions. These conditions are used by the speech



Fig. 2. Screenshot of a choice moment in the application

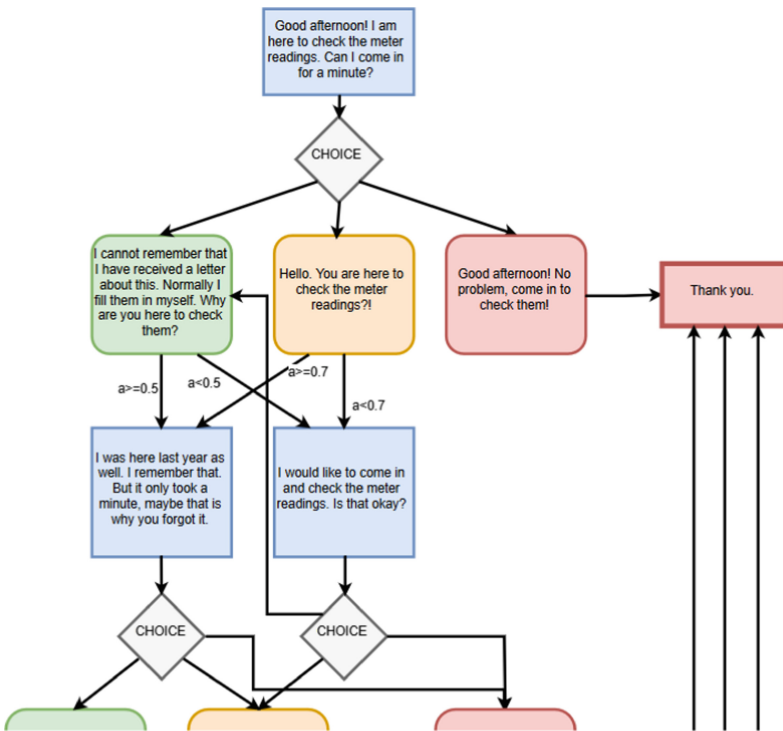


Fig. 3. Part of a (translated) flowchart for a scenario

analysis and indicate what the influence of the assertiveness of the players voice is on the progress of the scenario.

When the speech analysis module is turned on, a few more steps are taken before continuing with the scenario.⁵ Right after choosing a response (other than the repeat option) all the other responses disappear and a countdown is shown, after which the recording automatically starts. The player is instructed by the game to read their response aloud. The recording automatically stops after a certain time of silence. The reaction of the virtual opponent in the next turn is determined by the response the player made and the score of the speech analysis (more on this in Sect. 3.4). If the reaction of the virtual opponent is not an end dialogue, the player will again be able to choose from different reactions. If an end dialogue is played the outro of the scenario is played after that dialogue. This can either be closing the door (door scenarios), walking away (street scenarios), or a hung up sound (phone scenario).

When the scenario has ended the player is asked whether or not he/she would call the emergency number in such a situation, the rationale behind this is described in [8]. After this the player will receive feedback on their choices and see its score. This feedback is divided in three categories: general feedback, extended feedback and tips. For all feedback a ‘read aloud’ option is available. Section 3.3 discusses how the feedback and score are established.

The flow of the application is programmed in a general way, so that it is easy to add or change the content of the scenarios. The scenario specific content (avatars, dialogues, responses, tips, and feedback) are therefore stored in a database. By linking the code to the database the specific content is shown within the application. An overview of all the databases and the data stored in these databases can be found in Table 1.

3.2 Virtual Avatars and Environment

For each scenario at the doorstep and on the street, a male and female avatar are created, resulting in a set of eight different avatars. For the two scenarios at the front door the company uniform of two Dutch companies is manually recreated. For the scenarios on the street modern clothes are chosen for the avatars.

In order to animate the avatars iClone⁶ is used. Both default animations and manually created animations are used as body animations. For the face animations, Facial Motion Capture⁷ software, with a plug-in for iClone, is used. During two recording sessions a male and a female actor their face expressions were recorded while also recording the voice for the different dialogues. Their voices are also used for the scenarios that are not using a virtual agent (phone scenarios). After recording the facial animations, some further editing was needed in order to improve the animations. This was necessary because the results from the Facial Motion Capture were insufficiently realistic.

The three dimensional environment used for the application features a small part of a residential area and a partly decorated house. The environment is

⁵ Note: sometimes a response is an action instead of a verbal response, in these cases the speech analysis step is always skipped.

⁶ www.reallusion.com/iclone/.

⁷ www.facewaretech.com.

Table 1. Overview of all used databases

Database	Description
Avatar	This database contains the names of all the available avatars and links them to the right scene (location) and scenario number
Dialogue	This database contains all the turns of the virtual opponent. Storing: <ul style="list-style-type: none"> - Scene number, scenario number and the dialogue ID - Text of the dialogue - The tip variable IDs that is turned false in this dialogue (if applicable, more details on this in Sect. 3.3) - The type of end state (if applicable) - The body animation of the virtual avatar (if applicable)
Responses	This database contains the responses that are linked to the dialogues. Storing the following: <ul style="list-style-type: none"> - Scene number, scenario number and the ID of the dialogue that the response is corresponding with - Text of the response - The ID of the default next dialogue - The threshold for the speech analysis score (if applicable) - The ID of the next dialogue when the speech analysis score is below the defined threshold, and the ID of the next dialogue when the score is above the threshold (if applicable) - The action (animation) linked to the response (if applicable) - The type of response (2 = good, 1 = average, 0 = bad) - A boolean if there is speech analysis for this response (default = true)
Tip	This database contains the tip variable IDs (corresponding with the dialogue table), their name and text
Feedback	This database contains the different types of outcomes. The IDs correspond with the type of end state defined in the dialogue table. Furthermore, it contains the name and the feedback text of the outcomes

created in Unity, using various assets from the Unity Asset Store⁸. The virtual avatars, including their body and facial animations are exported to Unity and added to the environment.

3.3 Feedback and Score

After each scenario the player receives feedback. The general feedback, shown by default, tells the player whether or not he/she has become a doorstep scam. This feedback is therefore depending on the type of the end state of the

⁸ <https://assetstore.unity.com/>.

scenario, which is stored in the feedback database (see Table 1). Furthermore, the general feedback encourages the player to further explore the other feedback.

The extended feedback includes a paragraph that is specific for the scenario. It is therefore independent of the outcome reached by the player. Furthermore, the extended feedback includes feedback on the choice the player made for the call question. There are three different feedback texts available for this. The first is for players who chose to call the emergency number, the second is for people who did not choose to call the emergency number but did have a negative outcome, and the last is for people who did not call the emergency number but did also not become a victim in the scenario.

The most tailored feedback are the tips. In each scenario a number of variables are defined, who are by default true. If a specific dialogue that is linked (which is stored in the dialogue database) to a variable is played during the scenario, this variable is set to false. For the variables that are still true at the end of the scenario a tip is given to the player. The tips are showed in such an order that they follow the progress of the scenario. An example is a variable about asking the identification, this is linked to the dialogue where the virtual opponent shows its identification. If this dialogue is not played during the scenario (so the variable is still true and not set to false), the player receives feedback about this afterwards.

Besides feedback the player also receives a score at the end of each scenario. The higher the score of the player, the better the performance during the training was. The highest score is 105. The score is calculated using the average score for the choices, the average score for the voice, a score for the result of the scenario, and a bonus (of 5 points) if the emergency number is called.

Next to a score a player also gets a number of stars (0–5) for a scenario. A player will receive one star if he/she earned more than 11 points during the game, two stars for more than 33 points, three stars for more than 55 points, four stars for more than 77 points, and the maximal number of five stars is achieved when more than 100 points are earned during the game.

3.4 Speech Analysis

The speech analysis module is based on the Interpersonal Stances theory. This concept stems from social psychology, and can be defined as ‘the ways in which speakers and writers linguistically demonstrate their commitment to or attitudes about a person or proposition’ [6]. The module classifies the speech in 2 types of attitudes: Dominant (normally referred to as Above) and Submissive (Below).

A modified version of the openSmile⁹ toolkit extracts the voice features, while an SVM algorithm classifies the extracted features into the categories. The SVM model was built using 4-fold cross validation over a dataset with 681 sentences of four people instructed in how to act into both categories. Details about the algorithm and the SVM tuning are described in [5]. The final accuracy of the module is 86.56%.

⁹ <https://audeering.com/technology/opensmile/>.

The recorded voice of the player is analyzed by the algorithm. The application detects silence to determine when the recording ends. The ambient noise is measured while people are using the application. Silence is defined as a period of 3 s in which the volume is 20 decibels above the ambient noise volume maximum. The recording stops after 10 s or if silence is detected. The module runs on a server, it receives the user's voice and returns the classification status. In case of communication failure, the client application ignores the voice information and continues the dialogue without the speech analysis. The output of the module is the confidence percentage between the 2 categories. The application uses those values in the dialogue tree as described in Sect. 3.1. Figure 1 shows how the speech analysis is embedded in the flow of the application.

4 Evaluation Prototype

To obtain feedback from potential end users on the prototype, we organized a focus group at KBO-PCOB in Utrecht. This focus group consisted of five security advisors voluntarily working for KBO-PCOB. All participants belong to the group of potential end users, namely elderly. Two of the participants were female and three were male.

The focus group started with general instructions about the application. We explained how to start the application, introduced the different scenarios, and told the participants how to use the speech analysis module. After the instructions, all participants were provided a private practice space and a tablet. They each had 30 min to test the application and try different options. After this, we had a group discussion about their findings. This group discussion was based on open questions to get qualitative feedback.

The participants all responded positively about the application. They believed it is a nice addition to existing education on safety for elderly. They recognized the scenarios as being most common to experience in daily life. They especially saw added value of the application for group meetings in which elderly can work together on the application and afterwards discuss this with the group. They had some minor comments as well. For instance, they would like to see a progress bar in the scenarios with speech analysis so they know how much time they have left for recording their voice. Within those scenarios, they also wanted the time the system takes to detect the end of their spoken input to be shorter. Additionally, they would like to be able to go back within a scenario, to alter their answers, and to go back to the tips/feedback menu when clicking a scenario in the score screen. Finally, they would like to see more scenarios and they preferred lighter colors. Based on these suggestions, we are currently working on improving the application.

5 Conclusion and Discussion

In this paper, a serious game for verbal resilience training was presented. The training is specifically developed for elderly in order to teach them how to deal with scams. Elderly form a high-risk demographic group, and are relatively often confronted with scams, either on the phone, the street or at the front door. The training allows them to practice different scenarios on a tablet. Within the scenarios, the trainee is confronted with a virtual scam artist trying to convince the trainee to either pay money, provide personal information or let him/her inside the house. The trainee can interact with the virtual agent by choosing an option from the fixed multiple-choice menu and by either clicking on this answer or speaking the answer aloud. In the latter case, the speech recognition module is able to detect the level of assertiveness in the recorded voice, providing real-time feedback. Next to this, the trainees receive their score and feedback on their performance including tips at the end of a scenario and they can view the top ten scores and the average numbers of earned stars on a score page.

To evaluate the prototype we organized a focus group with potential end users. They reacted positively towards the developed environment. They especially saw the added value in the possibility to train scenarios in larger settings in which they can guide other users. They had some small remarks regarding the speech recognition module, which will be addressed in the final version.

References

1. Bosse, T., Gerritsen, C.: Towards serious gaming for communication training - a pilot study with police academy students. In: Poppe, R., Meyer, J.-J., Veltkamp, R., Dastani, M. (eds.) *INTETAIN 2016* 2016. *LNICSSITE*, vol. 178, pp. 13–22. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-49616-0_2
2. Bosse, T., Gerritsen, C., De Man, J.: An intelligent system for aggression de-escalation training. In: *ECAI*, pp. 1805–1811 (2016)
3. Chi, H., Agama, E., Prodanoff, Z.G.: Developing serious games to promote cognitive abilities for the elderly. In: *2017 IEEE 5th International Conference on Serious Games and Applications for Health, SeGAH 2017*, pp. 1–8 (2017)
4. Eisler, R.M., Miller, P.M., Hersen, M.: Components of assertive behavior. *J. Clin. Psychol.* **29**(3), 295–299 (1973)
5. Formolo, D., Bosse, T.: Towards interactive agents that infer emotions from voice and context information. *EAI Endorsed Transactions on Creative Technologies* **4**(10) (2017)
6. Gales, T.: Identifying interpersonal stance in threatening discourse: an appraisal analysis. *Discourse Stud.* **13**(1), 27–46 (2011)
7. Konstantinidis, E.I., Billis, A.S., Mouzakidis, C.A., Zilidou, V.I., Antoniou, P.E., Bamidis, P.D.: Design, implementation, and wide pilot deployment of FitForAll: an easy to use exergaming platform improving physical fitness and life quality of senior citizens. *IEEE J. Biomed. Health Inform.* **20**(1), 189–200 (2016)
8. van der Lubbe, L.M., Bosse, T., Gerritsen, C.: Design of an agent-based learning environment for high-risk doorstep scam victims. In: Bajo, J., et al. (eds.) *PAAMS 2018*. *CCIS*, vol. 887, pp. 335–347. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94779-2_29

9. McFall, R.M., Lillesand, D.B.: Behavior rehearsal with modeling and coaching in assertion training. *J. Abnorm. Psychol.* **77**(3), 313–23 (1971)
10. Miller-Day, M.A., Alberts, J., Hecht, M.L., Trost, M.R., Krizek, R.L.: *Adolescent Relationships and Drug Use*. Lawrence Erlbaum Associates, Mahwah (2014)
11. Nichols, T.R., Birnel, S., Graber, J.A., Brooks-Gunn, J., Botvin, G.J.: Refusal skill ability: an examination of adolescent perceptions of effectiveness. *J. Primary Prevent.* **31**(3), 127–37 (2010)
12. Nichols, T.R., Graber, J.A., Brooks-Gunn, J., Botvin, G.J.: Ways to say no: refusal skill strategies among urban adolescents. *Am. J. Health Behav.* **30**(3), 227–36 (2006)
13. Senger, J., et al.: Serious gaming: enhancing the quality of life among the elderly through play with the multimedia platform SilverGame. In: Wichert, R., Eberhardt, B. (eds.) *Ambient Assisted Living*. ATSC, pp. 317–331. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-27491-6_23
14. Taylor, L.M., Maddison, R., Pfaeffli, L.A., Rawstorn, J.C., Gant, N., Kerse, N.M.: Activity and energy expenditure in older people playing active video Games. *Arch. Phys. Med. Rehabil.* **93**(12), 2281–2286 (2012)