Multimodal Technologies and Interaction

MDPI

Article

Affective Communication between ECAs and Users in Collaborative Virtual Environments: The REVERIE European Parliament Use Case

Ioannis Doumanis [1,*] and Daphne Economou [2]

[1] School of Physical Sciences and Computing, University of Central Lancashire, Preston PR1 2HE, UK
[2] School of Computer Science and Engineering, University of Westminster, London W1B 2HW, UK; D.Economou@westminster.ac.uk
* Correspondence: idoumanis@uclan.ac.uk

Received: 31 December 2018; Accepted: 29 January 2019; Published: 10 February 2019

Abstract: This paper discusses the enactment and evaluation of Embodied Conversational Agents (ECA) capable of affective communication in Collaborative Virtual Environments (CVE) for learning. The CVE discussed is a reconstruction of the European Parliament in Brussels developed using the REVERIE (Real and Virtual Engagement In Realistic Immersive Environment) framework. REVERIE is a framework designed to support the creation of CVEs populated by ECAs capable of natural human-like behaviour, physical interaction and engagement. The ECA provides a tour of the virtual parliament and participates in the learning activity as an intervention mechanism to engage students. The ECA is capable of immediacy behaviour (verbal and non-verbal) and interactions to support a dialogic learning scenario. The design of the ECA is grounded on a theoretical framework that addresses the required characteristics of the ECA to successfully support collaborative learning. In this paper, we discuss the Heuristic Evaluation of the REVERIE ECA which revealed a wealth of usability problems that led to the development of a list of design recommendations to improve their usability, including its immediacy behaviours and interactions. An ECA capable of effectively creating rapport should result in more positive experiences for participants and better learning results for students in dialogic learning scenarios. Future work aims to evaluate this hypothesis in real-world scenarios with teachers and students participating in a shared virtual educational experience.

Keywords: Collaborative Virtual Environments; dialogic learning; Embodied Conversational Agents (ECAs); affective communication; REVERIE; rapport; immediacy behaviours

1. Introduction

Embodied Conversational Agents (ECAs) are autonomous software systems present in an interface with some form of embodiment. Their purpose is to mimic interpersonal communication when interacting with users, including the ability to produce and respond to verbal and nonverbal communication. In Virtual Environments (VEs), ECAs typically share the same virtual space with human users represented as avatars and have a role to play (e.g., tutors [1,2], virtual patients [3] and tour guides [4]). To make ECAs, more human-like and believable in their role researchers have used various computational approaches (e.g., personality and emotional models, advanced natural language processing, etc.). REVERIE (Real and Virtual Engagement In Realistic Immersive Environment) [5,6] is a framework which supports the creation of ECAs capable of recognising the user affect state and respond with appropriate verbal and nonverbal behaviours designed to facilitate social interaction among human users. We designed a collaborative learning scenario using the REVERIE framework. In this scenario, the ECA has the task of guiding groups of users (one teacher and six students) in

a virtual representation of the European Parliament in Brussels [6]. The ECA follows a narrative (e.g., with information about points of interest in the parliament) defined in its script. After the completion of the tour, students participate in an online debate on the topic of multiculturalism moderated by the teacher. The ECA does not have an active role in the debate. However, it keeps monitoring the users' activity, and interferes when needed (e.g., to call for attention when a student does not engage). The ECA is a female character capable of recognising the user's speech (yes/no keywords) and affect and to react accordingly using verbal and nonverbal immediacy behaviours.

The ECA's nonverbal behaviours are generated automatically based on a computational model [7] of socio-emotional attitudes. The computational model was developed following the Interpersonal Circumplex [8], which defines the personality space of the ECA based on the social attitudes of Dominant, Submissive, Friendly and Hostile [8] (see Figure 1). This model enables the ECA to use a wide range of non-verbal behaviours (including immediacy) as appropriate. The ECA's non-verbal behaviours and interactions with human users can be defined by an author in its script to match the requirements of a specific collaborative learning scenario. We believe that these bi-directional embodied interactions can be ascribed to the social communication patterns of attraction/repulsion and dominance/sub dominance [9]. Also, these interactions can be contained within the three dimensions (process, relationships, results) defined by the facilitation framework [10]. We have integrated these elements into a theoretical framework which provides a holistic view of the ECA's generated behaviours and its interactions with human users. Based on the theoretical framework we also propose a new model of rapport for ECAs. The model provides a first step for developing and validating ECAs that can effectively facilitate CL scenarios.

The ECA has full knowledge of the VE and can navigate through it autonomously. Users (represented as avatars) can follow the agent using mouse/keyboard/map navigation. The ECA can automatically engage the group to follow her guidance, but the user can still manipulate their avatar's position using one of the navigation methods mentioned above. We evaluated the ECA using a set of heuristics designed specifically for VEs [11]. The results showed several positive aspects of the current design of the ECA. However, they also showed 22 usability problems covering all aspects of the experience with the ECA (including its verbal and non-verbal behaviours and interactions). We have used these problems to develop design recommendations to improve the usability of the ECA. When implemented, these recommendations hold the potential to build ECAs that effectively create rapport with users in VEs. Rapport has been associated with effective learning in real-world and online learning scenarios [12,13]. It is possible to build rapport using traditional e-learning platforms (e.g., using emoticons [13] on Moodle) but it is a difficult and time-consuming process. Traditional e-learning platforms do not provide tools to gather information about the affective state of students either to facilitate the learning process. It is, therefore, reasonable to assume that an ECA integrating the aforementioned characteristics can have a positive impact on CL scenarios. We plan to evaluate this hypothesis with ECA displaying affective attitudes and capable of facilitating the learning process in dialogic learning scenarios.

The remaining of the paper is organized as follows: Section 2 explains the theoretical underpinnings of the ECA's affective behaviours; Section 3 gives a detailed account of the REVERIE modules which are involved in recognition of the user affect and the generation of the ECAs behaviours; Section 4 presents the evaluation of the virtual parliament (including the ECA) using heuristics; Section 5 presents the results of the study; Section 6 discusses the lessons learned from the evaluation; Section 7 provides an overview of future work, and the paper ends in Section 8 with the conclusions.

2. A Theoretical Framework of Affective Communication

For an ECA to help create rapport, it should exhibit immediacy behaviours and interactions which support collaborative learning. Immediacy behaviours are "approach behaviours" (verbal and non-verbal) conveying closeness between those communicating that increases stimulation of senses which is perceived as warmth [14]. Non-verbal immediacy can include behaviours such as moving

closer when you interact with someone, touching, using direct eye contact, smiling, having an open body posture, posture and vocal expressiveness. Verbal immediacy behaviours include the use of personal names, pronouns such as "you" and "we" and verbal empathy. A REVERIE ECA can be programmed to exhibit immediacy behaviours using a specific use case script. A script describes how the ECA should behave in specific situations (e.g., how to behave when students do not engage) within a specific scenario. A REVERIE ECA can exhibit a wide range of non-verbal behaviours (including immediacy) following the social-emotional attitudes of Dominant, Submissive, Hostile and Friendly [7,8] (see Figure 1). Non-verbally, there is a continuum between agency and communion [8] depending on the requirements of each situation.

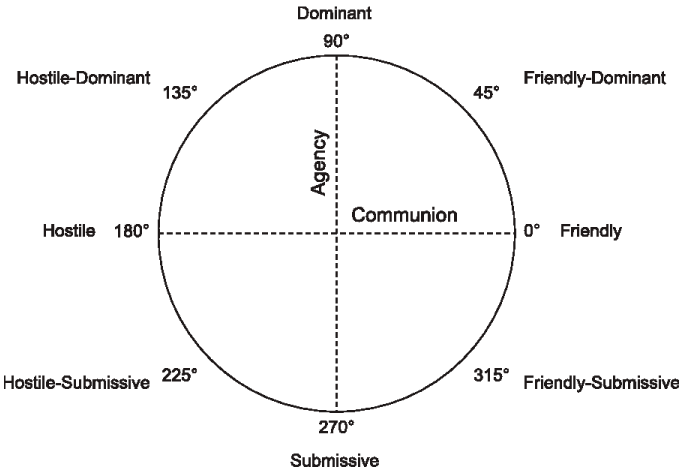

Figure 1. The Interpersonal Circumplex.

The ECA's script can also specify how it should interact with human users. For example, the ECA can analyse the students' engagement in a learning activity and call for attention to students who do not engage. Students can adhere to the ECA's request and become more engaged (as denoted by eye tracking and the user's head orientation) or ignore it. The ECA can "learn" on user behaviours and in time ignore users who do not frequently engage. This kind of embodied interaction between the ECA and users requires its theoretical underpinnings to describe properly. The theory of affective communication [9] postulates that human interactions are processed through the agent's corporeal body, and they can be described based on two dimensions of the felt body: (1) attraction and repulsion and (2) dominance and sub-dominance. The dimensions of attraction and repulsion determine whether the ECA and a user attractively converge or repulsively diverge in their perspectives during an online learning activity. These dimensions determine whether a user shows a tendency towards the ECA and the learning activity (attraction) which translates to engagement or they show a tendency away from the ECA and the learning activity (repulsion) which translates to non-engagement. Regarding the dimensions of dominance and sub-dominance, it is assumed that there is a dominant and subdominant side in ECA to human interactions, although sides may frequently change between the communicating partners. In the case of attraction, the dominant side becomes the perspective giver and the subdominant side the perspective taker. Specifically, the teacher (*dominant*) initially controls the ECA (*subdominant*) and decides when the tour should start. Then, the ECA takes over (*dominant*) and autonomously guides the group (*subdominant*) in the VE. It can also call for attention (*dominant*) to students (*subdominant*) who do not engage in the educational activity. Assuming student attraction as a pre-condition for regaining engagement and contingent behaviours between the ECA and the student (e.g., behavioural mimicry [15]), it may transfer its perspective and successfully re-engage students with itself and the educational experience.

We designed a range of immediacy behaviours (verbal and non-verbal) and embodied interactions for the ECA to support collaborative learning scenarios in REVERIE virtual parliament. This type of

learning involves groups of users learning or attempting to learn something together. The following table explains some of the characteristics of collaborative learning (see Table 1) [16].

Table 1. Characteristics of Collaborative Learning.

Characteristics	Collaborative Learning
Goal Structure	Collaborative
Role of Students	Active participation, autonomous learners
Role of teacher	Guide
Material used	Materials are arranged according to the purpose of learning
Types of activities	Various types of activities to engage learners in a shared learning community
Types of interaction	Intense student-student interaction
Classroom physical setup	CL groups
Teacher-Student Relationship	Collaborative and equal
Independence	None or negative
Learning expectations	Group success as well as individual's

In collaborative learning scenarios, the ECA should semi-autonomously (controlled by the teacher) or autonomously support teachers facilitating the learning process. In such scenarios, the ECA helps human participants (teachers and students) to move through the learning process together. Specifically, the ECA guides the group to share ideas, opinions, experience and expertise to achieve the learning goals of the sessions. For the ECA to successfully facilitate a collaborative learning process it needs to balance the focus across the following dimensions (see Figure 2) [10]:

- ***Process:*** There are many different types and approaches to collaborative learning. We have focused our investigations on the learning processes of dialogic learning [17]. In this type of collaborative learning, students deepen their understanding of a topic through listening, sharing and questioning. To ensure the ECA can assist teachers in managing a dialogic learning process it needs to display at minimum a range of basic facilitation skills.
- ***Relationships:*** To ensure the process works successfully and group members are engaged the ECA needs to manage affective relationships (see discussion above about the dimensions of attraction/repulsion dominance/sub-dominance). These refer to the relationship of the ECA with the group and how it helps build relationships between the group members. For the ECA to develop and maintain these relationships, it needs to be able to develop a rapport with all participants in the learning experience.
- ***Results:*** The process is result driven. This means that the ECA should assist the teacher to get to this destination during a dialogic learning session. The ECA can monitor the quality of the arguments used by students [18]. At the end of the session, it can assist teachers in providing tailored feedback to students by providing their performance data (e.g., where they did well and how they can improve)

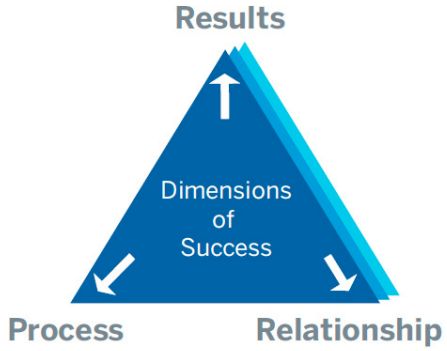

Figure 2. Facilitation Framework for Collaborative Learning.

The basic facilitation skills the ECA needs to exhibit to assist teachers in managing the process of dialogic collaborative learning are [19,20]:

- Making everyone feel comfortable and valued

The ECA needs to address participants by name and thank them for attending the learning session. It should also be able to check that they understand the goal of the session and its learning objectives.

- Increasing understanding

The ECA should be skilled at contributing relevant facts, stories and multimedia material (e.g., pics and videos) at the right point of the discussion to help increase the group's understanding of the topic.

- Encouraging participation

The ECA should encourage all group members to participate in the session. Some effective techniques to use are various types of questions (e.g., open-ended and probing questions) and verbal re-enforcers to reward the desired behaviour.

- Listening and observing

The ECA needs to be able to listen actively and hear what each group member is saying. It needs to be able to check the understanding of each participant (e.g., quality of the arguments contributed to the discussion) to ensure the learning objectives of the sessions will be met.

- Guiding the group

The ECA must be able to guide the group by managing basic logistics (e.g., time-keeping and referring to ground rules) and giving feedback (see the "*Results*" dimension above).

The above skills are provided only for guidance to assist in the design of behavioural models for ECAs. Careful empirical work is needed to validate properly and (re)define these skills.

At the minimum level, the ECA should behave as an autonomous intervention mechanism to ensure the active participation of students in the experience. Overall, the fusion of these elements creates a theoretical framework for affective communication for this paper. This framework covers the ECA's generated verbal and non-verbal behaviours and its embodied interactions with human users.

3. The Rapport Model

Based on the theoretical framework above, we designed the rapport model for ECAs. The model is based on two main domains represented by a horizontal and a vertical axis of Figure 3. For the horizontal axes, we designed a domain related to immediacy behaviour, which ranges from low to strong. The vertical axes represent two correlated variables related to the use of the ECA: "*interaction*" and "*Intervention/Facilitation*". Finally, we defined three levels in which the potential interaction of the ECA and its immediacy behaviours increase. These are:

1. At the engagement-driven level, the ECA should work as an automated intervention mechanism to ensure the student's engagement is maintained to the learning activity. The ECA should exhibit a range of immediacy behaviours (verbal and non-verbal) only in response to the student's engagement.
2. At the teacher-driven level, the ECA should be more participative in the learning process, implementing a range of immediacy behaviours (verbal and non-verbal), and supporting the learning process. The teacher should have control over the ECA's participation, i.e., when and how to support the learning process). The ECA should display some basic facilitator skills (e.g., increasing understanding) depending on the requirements of the scenario.

3. At the full facilitator level, the ECA should fully autonomously facilitate the learning process when asked by the teacher. It should display a wealth of facilitator skills (including basic facilitator skills) depending on the requirements of the scenario.

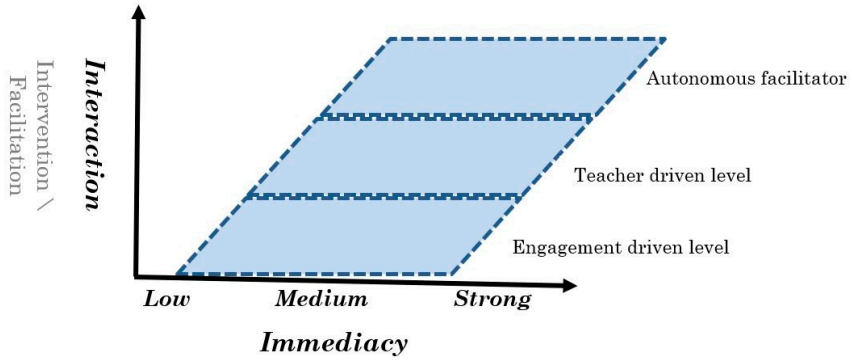

Figure 3. Proposed rapport model for an ECA to support collaborative learning.

4. The ECA of the REVERIE Virtual Parliament

This use case is about an educational excursion to a virtual representation of the European Parliament. It involves a group of users (a teacher and a student) and the ECA which has the role of the tour guide. The ECA fully implements the engagement-driven level of the rapport model (see Figure 4). When users' login to the system, they are greeted by the ECA who asks if they are ready to start the tour. Only the teacher can answer positively or negatively (yes/no) to the ECA's request. Once the tour starts, the ECA navigates users in predefined areas of the virtual parliament. During the tour users automatically follow the ECA, though they can choose to walk away from the group using a map-based navigation system. When the tour ended the ECA asked the group if they enjoyed the tour and invited them to participate in a debate. The debate is on the topic of multiculturalism which fully implements the characteristics of collaborative learning discussed in the theoretical framework (see Table 1). The ECA explains the process of the debate and moves to the side of the virtual scene. For the duration of the debate, the ECA monitors the engagement of the students and calls for attention to those who do not engage. The level of engagement of each user is also realised on the graphical user interface (GUI) using smileys (happy, neutral and disengaged) (see Figure 4) and shared between all participants of the experience.

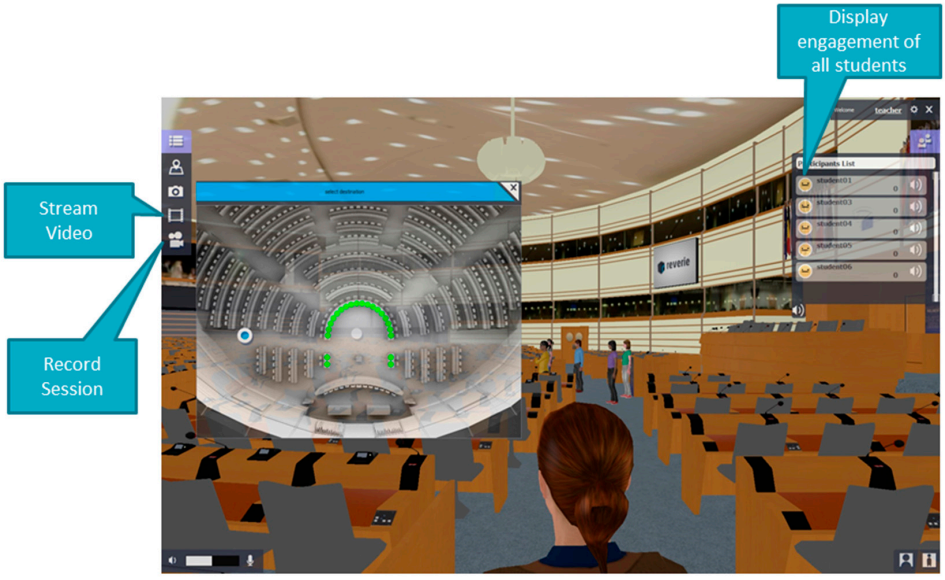

Figure 4. Level of engagement displayed on the GUI.

To realise this behaviour, the ECA uses four modules of the REVERIE platform. These modules enable the ECA to recognise, process and respond to multimodal input from users. They are explained in detail in the following sections.

4.1. Human Affect Analysis Module

The human affect analysis module consists of three components [21]:

1. ***User affect recognition:*** The user affect component provides a continuous prediction of the user's arousal and valence level using a standard web camera. Detecting key facial landmarks is an important step in recognising the user affect. The component can track and use 46 landmarks of the face as input to process arousal and valence levels.
2. ***Head Node and Shake Detection:*** This component identifies head gestures as an indication of agreement or disagreement. To detect head nodes and shakes the component measures differences in the direction of the head movement between two adjacent frames for a certain time.
3. ***Gaze Direction:*** This component identifies where the user is looking at and uses it as an indication of the level of engagement. Based on the user's head pose (position and orientation) the component determines whether the user is looking at the screen to determine engagement (see Figure 5). The method measures the head pose within a threshold. For as long as the user's head pose varies within the range, the user is considered engaged.

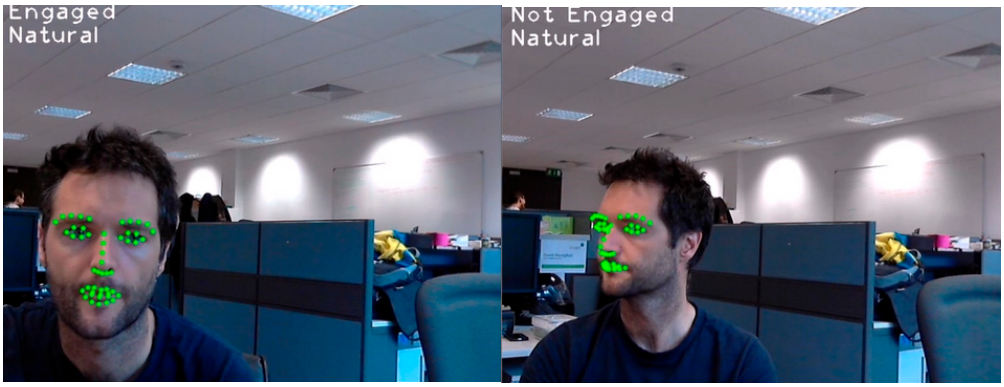

Figure 5. A user engaged with the ECA (**left side**) and not engaged with the ECA (**right side**).

4.2. Speech Analysis

The role of this module is to provide the ECA with the ability to recognise verbal input from the user. To meet the requirements of the educational scenario, the module recognises simple keywords ("yes", "no" and numbers). The speech analysis module uses the open source CMY Sphinx library [22] for speech recognition.

4.3. The Reasoning Module

The Reasoning Framework controls the ECA's behaviours. This module enables a human author to control the behaviour of an ECA using a use case specific script. The script specifies how the ECA should behave, what to say when and where, where to go, etc. It also offers several other options such as to specify additional information about the VE (e.g., points of interest), flow control to specify a non-linear story, context-specific behaviours (e.g., how to react when users are not engaged, etc.). An ECA can have a full multimodal interaction with a user or with a group of users. For both types, both spoken input, and user head movements (nod and shake) can be accepted as input. The *Speech Recognition* module supplies information on the spoken input and the *User Affect* analysis module supplies head nods and head shakes input. When the ECA addresses one user, it will recognise the first response as a valid answer. When it addresses a group, it will wait for the responses of all group

members; it will interpret the valid responses and deduce the answer of the group as a whole based on a simple majority rule [5,21].

4.4. Autonomous ECA Behaviours

In the current implementation, the ECA acts as an intervention mechanism when students don't engage in the learning experience [21]. The ECA exhibits mainly immediacy behaviours (verbal and non-verbal) and interactions. However, it may also exhibit some dominant but still friendly behaviours (e.g., eye gaze), especially when calling a student for attention.

Eye Contact

The ECA knows the current position of all users (represented as avatars) in the VE. When it addresses the group, it will look at individual members alternatively, sometimes for a short duration while others for a bit longer. The ECA will ignore users located behind its back. It will also maintain this gaze behaviour when giving a tour of the VE.

Proximity

At the beginning of the tour, the ECA can automatically engage participants in the tour using the Follow-Me mode. In this mode, the destinations of individual avatars are assigned by the system based on the destination and orientation of the ECA. This means that users will remain near the ECA throughout the tour. When the group reaches a destination in the VE, each participant is positioned in front of the ECA based on engagement and available space (see Figure 6).

Gesticulation and Facial Expressions

During the tour, the ECA can use a range of hand-gestures (e.g., pointing, beckoning) to accompany its speech in a friendly manner. It also displays three basic facial expressions (smile, frown, neutral) in synchrony with its speech.

Figure 6. Proximity of users with the ECA.

Verbal Immediacy

The agent uses speech that shows openness and empathy. For example, it uses pronouns such as "you" and "we". It also uses informal messages that encourage students to connect with her (e.g., *"Great I am glad you enjoyed the tour"* or *"So, what are we waiting for? Let's get on"*).

Interactive Behaviours

The ECA analyses the user engagement based on information it receives about the user's attention from the "User Affect Analysis" module. To detect the user's engagement, the ECA tracks the user's eyes and head orientation. Based on the analysis of these characteristics it determines a level of

engagement. If a user is not engaged, it will show on their avatar's gaze (their avatar seems to look over their shoulder as shown in Figure 7b below), and it will trigger the ECA's reaction. In the current study, the ECA walks to the user (but not too close), makes eye contact, calls for attention and walks back to its original position. If a user does not frequently engage, in time the ECA will learn to ignore them.

(a) (b)

Figure 7. (**a**) A user engaged with the ECA (right side) and (**b**) the same user not engaged with the ECA (left side).

5. Heuristic Evaluation of the REVERIE ECA

We evaluated the ECA as part of the overall usability evaluation of the REVERIE VP use case. We decided to use heuristic evaluation for two reasons: (a) we had already conducted user testing with previous versions of the REVERIE VP prototype, and (b) budget and time constraints prevented us from running more testing with real users. We invited three evaluators and asked them to review the prototype using a list of heuristics specifically designed for evaluating virtual reality environments [23]. The number of evaluators recommended in any heuristic evaluation is three to five since there is not as much additional feedback to get from larger numbers [24]. The expert evaluators were given a set of tasks to complete with the prototype and a list of heuristics to evaluate each task against it. They all worked at the same time in a room interacting in the virtual environment using different computers. During the session, they did not communicate with each other. After experts had completed the tasks, they participated in a debriefing session moderated by a REVERIE researcher to have their findings aggregated.

5.1. The Evaluators

The group of evaluators were chosen to include persons with expertise in UX and video games (see Table 2). The gamer evaluator was selected to provide reactions representative to the intended user group. Each evaluator conducted the heuristic evaluation of the prototype (including the ECA) individually. The table below summarised the profiles of each evaluator.

Some tasks required evaluators to communicate with each other, but they were advised not to share any of the usability problems they found. This was important to ensure an independent and unbiased evaluation for each evaluator.

Table 2. The profiles of the three evaluators.

Evaluator	Profile	Gender	Age Group
Evaluator 1 (Teacher)	Senior Lecturer in Computer Science	F	35–45
Evaluator 2 (Student1)	Senior Lecturer in Computer Games Development	M	35–45
Evaluators 3 (Student2)	Video games player	M	25

5.2. Heuristic Evaluation

Literature suggests several sets of heuristics that can be used in the evaluation of virtual environments like the REVERIE VP. We selected the following set of heuristics (see Table 3) because it has been empirically validated. Specifically, its developers have shown that the particular set of heuristics can identify usability problems in virtual environments that cannot be found using traditional heuristics (e.g., Nielsen's Heuristics [25]). The set includes 16 heuristics which can be grouped into three categories: [11] (a) Design and aesthetics which includes the first four heuristics (H1–H4); (b) Control and Navigation which includes the next eight heuristics (H5–H12) and (c) Errors and Help which includes the last four heuristics (H13–H16).

Table 3. The list of 16 heuristics used in the evaluation of the prototype.

ID	Definition	Explanation
H1	Feedback	A VW must always keep the user informed about the condition of its avatar, related events or relevant facts that occur inside the VW. The virtual world must provide feedback easily noticeable to the user in view of any action that he begins or that can affect directly or indirectly.
H2	Clarity	A VW must have a control panel which is easy for the user to understand, using a clear language. Also, the elements of the control panel should be shown tidy and put into groups in a way that the user can be able to find what he is looking for intuitively.
H3	Consistency	The VW must be consistent in all its aspects, this way the user can predict the result of every fulfilled action.
H4	Simplicity	The control panel of the VW should not be overloaded, and it must have only the needed and relevant information. The icons, the messages of the system and the interaction with the objects inside the VW must be simple and intuitive.
H5	Orientation and navigation	A VW must have intuitive navigation and memorable. And must give the user a way to locate it inside and a way to find a determined location.
H6	Camera control and visualization	A VW must allow the user to determine the level and quality of the texture, visual effects or objects which purpose is only aesthetic. Also, the VW must give the user the control of the camera or angle from where it is visualised.
H7	Low memory load:	A VW must minimise the demands on the user's memory by making the objects, options and actions visible and easily accessible. Also, the system must provide the user ways to register or remember places inside the VW that have been visited or that can be of one's interest.
H8	Avatar's customization:	A VW must offer a whole group of predefined avatars, with a specific genre, age, appearance, among other attributes. The VW must allow the user to change the aspects of the avatar, whenever he wants.
H9	Flexibility and efficiency of use:	A VW must provide accelerators for common actions and allow the user to define his accelerators and change other interface options. This allows the advanced users to interact with the VW more efficiently.
H10	Communication between avatars:	The interaction inside the VW must be analogous to the real world. It must be done easily and intuitively and should be clear for the emitter and the receptor.
H11	Sense of ownership:	The physical rules of the real world should be maintained in a VW. In the case that the rules could be changed; the VW must inform these variations in a clear and explicit way.
H12	Interaction with the Virtual World:	A VW must indicate to the users which objects of the VW they can interact with and the ones they can't, also indicating which actions can be carried out with the objects that they can interact.
H13	Support to learning:	The complex objects of a VW must be complemented with definitions and indications for its use; this way learning may be promoted.
H14	Error prevention:	The VW must prevent users from making any mistake or creating undesired situations, related to the interface or the VW.
H15	Helps users to recover from errors	A VW must provide the user with the tools to recover from system errors or any undesired situation when the user cannot recover by himself.
H16	Help and documentation	A VW must give the user relevant information not only online but also inside the VW. This information must be of easy access and written or spoken in the user's language.

5.3. Evaluation Procedure

Evaluators were asked to review the prototype twice. The goal of the first review was to get a feel of the flow of the interaction and the general scope of the use case scenario. Evaluators were given a printed GUI map and were asked not to record any usability problems. Once the evaluators indicated they were ready, they were asked to start the second review. A moderator provided a copy of the list

of heuristics to each evaluator and explained the process of heuristic evaluation. The evaluators had to explain using a notepad whether the design of the prototype matches or violates a heuristic. Similarly, if an evaluator thought that the design of the prototype violates a heuristic, they had to explain why using the same notepad. The result was a list of usability problems specified by each evaluator and the heuristic each problem violates. At the end of the session, evaluators participated in a moderated debriefing session to have their findings aggregated. The whole session was recorded in audio and later transcribed to assist in the review of the aggregated results. We used QSR NVivo 12 to organise, search, code and evaluate the data.

5.4. Results

Table 4 shows the number of comments evaluators made about the ECA. The three evaluators made a total of 88 comments covering both the immediacy behaviours (verbal and non-verbal) and interactions of the ECA.

Table 4. Evaluator Comments for the ECA.

Cluster	Number of Comments
Non-verbal Immediacy	10
Verbal Immediacy	3
Immediacy Interactions	75
Total	88

We converted these comments into 22 usability problems and presented our findings below. For each usability problem (UP) identified we also indicate the heuristic it violates.

5.4.1. Verbal and Non-verbal immediacy behaviours

In this cluster, we looked for usability problems related to the ECA's verbal and non-verbal immediacy behaviours.

- *UP1: Absence of instruction on how to use REVERIE limited the ECA's helpfulness (H1)*

The evaluators thought the ECA presented information during the tour in a clear and consistent manner. However, they also thought that the absence of instructions on how to use REVERIE (e.g., UI and various tools) limited its helpfulness.

- *UP2: The ECA's voice is too synthetic (H2)*

As for the ECA's voice, the evaluators thought it was clear without up/down fluctuations which facilitated recognition by participants. However, the voice was too synthetic which limited the ECA's realism and impact.

- *UP3: Limited realism of the ECA's verbal behaviours (H13)*

The evaluators found the acoustic reverberation (3D spatial audio) when the ECA spoke realistic. However, there were issues with the 3D spatial audio engine which limited the impact. For example, evaluators commented that when the ECA stops talking the spatial audio comes to an abrupt stop which creates an awkward silence (absence of environmental ambience) that spoils realism. Other comments include listening to random steps echoing in the VE during the experience.

- *UP4: Low graphical realism of the ECA (H6)*

The evaluators agreed that the graphical realism of the ECA was low (e.g., texturing and lighting). Although it did not render the ECA annoying, it made it difficult for evaluators to connect at any level.

- *UP5: The ECA does not orient itself correctly at the beginning of the tour (H10)*

The ECA maintained its gaze with all users throughout the tour. However, the evaluators thought it was frustrating that the ECA did not orient itself towards the group at the beginning of the tour.

- *UP6: Follow-me does work for all users (H5)*

All evaluators agreed that Follow-me is a useful feature to have, especially in online classroom scenarios. However, it did not properly work for all users. For one evaluator the feature worked from the middle of the tour and not from the beginning.

5.4.2. Interactive Behaviours

In this cluster, we looked for usability problems related to five types of interactions: ECA—group (teacher and students), ECA—teacher, ECA—students, and teacher—students.

 # ECA—group (teacher and students)

- *UP7: Absence of empathy at the beginning of the tour (H7)*

The evaluators agreed that the *"Follow-me"* option is useful in online classroom scenarios. However, to make the ECA appear more empathetic, it should ask users to follow (e.g., *"Would you like to follow me during the tour?"*) before using *"Follow-me"*.

- *UP8: The tour requires users (teachers and students) to split their attention between information about multiple points of interest in the VE (H7):*

The tour requires users to split their attention [26] between, and mentally integrate, multiple sources of information (e.g., information about two consecutive locations).

- *UP9: Absence of control over spatial 3D audio (H2)*

The spatial 3D audio enhanced the realism of the experience. However, evaluators expressed concerns that using this type of audio with a large group of students would produce noise (e.g., when students talk on top of each other).

 # ECA—teacher

- *UP10: Multimodal input to the ECA does not always work (H10):*

The ECA does not always respond to multimodal input (e.g., it did not respond to the teacher's nod at the beginning of the tour).

- *UP11: Cannot distinguish between students who appear disengaged and those who are disengaged*

The evaluators perceived positively the ECA's ability to recognise the students' engagement. However, they expressed concerns about whether the ECA could distinguish between a genuinely disengaged student and a student who appears to be disengaged (e.g., someone who takes notes and looks down)

- *UP12: Low usefulness of the ECA in scenarios with a low number of students (H4)*

The evaluators agreed that the ECA's presence was not necessary in this scenario. This is because it was relatively easy for the teacher to monitor students and act accordingly if they do not engage. The ECA would be more useful in educational scenarios where a teacher is required to manage large groups (20–30) of students online.

- *UP13: Absence of teacher control over the ECA (H9)*

The ECA's autonomous behaviours can disrupt a learning session (e.g., in case of an error in engagement detection). The evaluators agreed that the ECA should act in a semi-autonomous manner controlled by the teacher.

 # ECA—Students

- *UP14: Lack of feedback about the ECA's internal state (H1)*

As the ECA did not provide feedback about its internal state (e.g., why it thought a student did not engage), evaluators had a hard time understanding what triggered its reactions.

- *UP15: Unrealistic and repetitive behaviours towards unengaged students (H15)*

The evaluators agreed that the behaviours the ECA conveyed to the students who did not engage in the educational activity were unrealistic and repetitive. This refers to the design of both verbal and non-verbal behaviours of the ECA.

- *UP16: Unrealistic and repetitive behaviours towards students who appeared unengaged. (H15)*

An ECA approaching students who appear unengaged with unrealistic and repetitive behaviours can distract the learning activity as a whole.

- *UP17: Slow update of affect states on the GUI (H16)*

There was an observable latency (5–6 s) between affect detection for a user and the update of the UI of all users.

- *UP18: Lack of natural user responses (H10)*

The evaluators did not always respond naturally during the learning session. They exhibited a wide range of deliberate behaviours with the goal of testing the affect recognition of the ECA.

 # Teacher—Students

- *UP19: Teacher engagement with unengaged students (assuming large classes) (H10)*

Even if the teacher had reliable information about the engagement of students, it would be difficult to react with a large group of students. The evaluators agreed that a method is needed for teachers to engage unengaged students without disrupting the rest of the group.

- *UP20: Lack of deep engagement of students with an educational task (H13)*

Evaluators sought the need for educational activities that engage students deeper in a learning scenario. For example, a student can check the microphone of the lectern (e.g., 1–2 testing) before starting their presentation.

- *UP21: Difficulty of keeping the turn-taking of students online (H13)*

Turn-taking can work effectively if students are cooperating. However, it was difficult for evaluators to find an alternative if students are not cooperating.

- *UP22: The debate requires teachers to split their attention between multiple student presentations (H7)*

The debate requires teachers to split their attention between multiple presentations to evaluate a student's performance.

5.4.3. Nature of the Problems

We took a closer look at the usability problems above with the goal to identify more information about the nature of the problems. Figure 8, shows the categorisation of the identified heuristic violations in the REVERIE ECA. As shown in the figure from the 16 usability heuristics we used, communication between avatars, support to learning and low memory load were the three heuristics most frequently violated by the current design of the ECA. The three categories alone accounted for 90.9% of the identified violations.

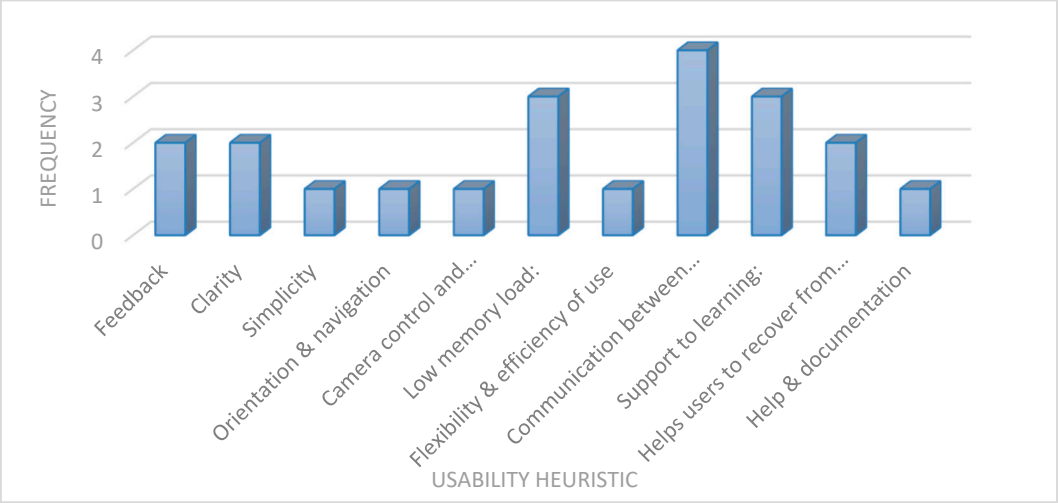

Figure 8. Categorisation of heuristic violations for the REVERIE ECA.

5.5. Discussion

Beginning the discussion of the results with the ECA's immediacy behaviours (verbal and non-verbal) it is clear that the ECA's verbal immediacy behaviours worked to a degree. The immediacy attributes implemented in the ECA's script (e.g., use of pronouns and informal manner of addressing users) had a positive impact on how evaluators perceived the clarity and consistency of its verbal behaviours. However, the evaluators felt that: (a) the artificial voice of the ECA and (b) the absence of instruction on how to use REVERIE limited its helpfulness. Additional issues with the use of 3D spatial audio further limited the impact of the ECA's verbal behaviours. Then, the evaluators found the non-verbal immediacy behaviours of the ECA to be the most problematic. They thought that the low graphical realism of the ECA and aspects of its non-verbal behaviours (e.g., lack of proper orientation towards the group at the beginning of the tour) convened little immediacy.

About the way the ECA interacts with users, it is clear from the results that the teacher should be in a *dominant* position when interacting with the ECA. The ECA should not be able to interfere (*sub-dominant*) in the learning activity unless the teacher explicitly requests it. As the teacher has input about the affective state of students through multiple modalities (GUI icons and avatars), it is relatively straightforward to know when the ECA is needed. Another important finding is that the ECA cannot distinguish between the states of *"disengaged"* and *"appears disengaged"*. As the gaze direction component does not have a *"not tracking"* state, it is difficult for the ECA to know when students are not looking at the screen (e.g., because they are taking notes) to categorise them as appropriate. Then, it was difficult for evaluators to immediately understand which of their actions (e.g., looking down or turning their head around) triggered the ECA to call for attention. They gave two key reasons for the lack of immediate understanding. First, it takes longer to update the user's affect state on the GUI (e.g., to change happy to neutral). Second, the ECA does not provide any feedback to explain what triggered its reaction. An important finding related to sharing of affect information on the GUI is that it may facilitate unnatural student behaviours. It is possible that students may not react naturally knowing they are being watched by an ECA which monitors their emotional and engagement status

and shares the information with teachers. Evaluators thought that the behaviours the ECA exhibited when it called for attention were unnatural and repetitive. Using such an ECA to (re) engage students who appear disengaged can disrupt a learning experience as a whole. The evaluators also sought the importance of learning content that eliminates the need to integrate multiple sources of information and deeply engage students in the learning scenario. Finally, the evaluators discovered problems related to managing large groups of students on the REVERIE prototype (e.g., the lack of tools to maintain turn taking).

Overall, the ECA violated a certain number of usability problems, including communication between avatars, support to learning and low memory load. While these usability violations are specific to the REVERIE ECA, we believe they reflect the usability problems in the design of the specific type of ECAs in general. As technology makes the ECAs smarter, their utility in collaborative learning scenarios will depend on how teachers can use the ECA and the behaviours (verbal and non-verbal) it exhibits. In the next section, we provide a series of design recommendations on how to address the problems we identified. When implemented, these recommendations hold the potential to disrupt all levels of the proposed rapport model (see Figure 2). Creating rapport with students has been associated with effective learning in both real life [12] and online learning scenarios [13]. For example, it has been shown that relatively simple strategies implemented using standard Web technologies (e.g., emoticons to create a supportive tone) can increase teacher-student immediacy online [13]. It is possible, therefore, that an ECA built following our recommendations can positively impact the students' learning performance in collaborative dialogic learning scenarios.

6. Design Recommendations

Based on the findings of the study reported above, we created a list of design recommendations for optimising the usability of the ECA when used in collaborative dialogic learning scenarios. To prioritise the recommendations, we used a custom nine-point scale (0 = not important, 8 = extremely important) inspired by the planning poker agile method [27]. These recommendations are highly actionable and situated to ECAs and hence, can be implemented in similar ECAs. As for the REVERIE ECA, we deemed these recommendations necessary for future versions of the prototype. In total we have identified ten important recommendations that we present in layman terms below:

1. Avoid displaying affect information on the GUI (priority = 8)

Providing affect information to students can create problems in a dialogic learning scenario (e.g., students feeling they are constantly watched). Do not include affect information on the GUI to ensure students respond naturally throughout the learning experience.

2. Improve the verbal immediacy cues of the ECA (priority = 8)

Consider the following improvements in the linguistic cues of the ECA:

- greet each user (teachers and students) by name;
- add verbal empathy (e.g., ask users to follow the ECA before using Follow-me);
- blend the ECA's speech with natural sounds in the VE (e.g., environmental ambience can fade-away when the ECA speaks).

The ECA should display the above linguistic cues using a human-like voice.

3. Improve learning content (priority = 8)

This recommendation is about improving the learning content to further support the ECA's verbal immediacy. Consider making the following improvements:

- add an introductory section at the beginning of the tour with instructions about how to use REVERIE (UI and multimedia/multimodal tools);

- eliminate the need to mentally integrate multiple sources of information by providing an option to repeat content (in a paraphrased/similar manner) about (a) Points of interest in the VE. (b) Presentations of students during the educational activity.

4. Improve the non-verbal immediacy cues of the ECA (priority = 8)

Consider the following improvements in the non-verbal immediacy cues of the ECA:

- orient the ECA towards the bulk of the users at the beginning of the tour;
- constantly monitoring the proximity of the users to ensure they follow the group.

5. Enable teacher-driven participation of the ECA in the learning process (priority = 8)

Consider choosing between the following interaction schemas to provide teachers with the necessary control over the ECA.

1. As teachers can receive information about the students' affective state through multiple modalities (avatar and GUI (in the form of smileys)), they should be able to ask the ECA for help when they need it the most.
2. The ECA could alert the teacher when a student is disengaged - from there the teacher can decide if the ECA should engage the student or if the teacher should handle the situation in real life.

In both schemas, the UI should allow teachers to target the unengaged students (e.g., by selecting their names on the GUI) separately from the rest of the class.

6. Add a "no tracking" state in the gaze direction component (priority = 7)

Add a "*no tracking*" state to the gaze direction component to signify when users are not looking at the screen. Setting a time threshold for how long the user can look away from the screen will enable the ECA to know when a student appears disengaged.

7. Allow for interpretation of the ECA's internal state (priority = 8)

8. The current design does not allow users to interpret what triggers the ECA's reaction when it calls for attention. To aid understanding, consider providing students with a snapshot of their engagement and emotional status (e.g., head orientation and emotions exhibited) on the UI.Build crowdsourced engagement behaviours for the ECA (priority = 8)

In the real-world teachers use various ways to get and keep their student's attention in the class [28]. Enabling teachers to build their own engagement behaviours (verbal and non-verbal) will endow the ECA with a large pool of behaviours to use with students. In turn, this should make the ECA appear more natural and versatile in its responses to students.

9. Use the ECA during the educational experience to encourage student participation (priority = 7)

The ECA can autonomously (or semi-autonomously) participate in the learning experience to ensure deep student engagement. For example, when watching a presentation by one of their peers, the ECA may motivate students to record a video or to take notes to use in group discussions.

10. The ECA should assist teachers in guiding the group (priority = 7)

The ECA should be able to assist teachers in keeping the ground rules for an effective learning session. At the basic level, these rules are time-keeping and turn-taking between students.

7. Future Work

Future work will focus on two areas. First, we will update the ECA's based on the design recommendations generated from the study and use it in a teacher-driven dialogic learning scenario. This means that the ECA should assist teachers in facilitating the learning process in a semi-autonomous manner (see level 2 of Figure 2). It should also exhibit some of the standard facilitator skills discussed in the theoretical framework. Second, we would like to evaluate the hypothesis about such an ECA

having a positive impact on the learning performance of students in dialogic learning scenarios. We plan to run a study to compare the impact of different kinds of ECAs (e.g., an ECA implementing our recommended behaviours vs an ECA who does not pay attention to students) on the performance of students. We hope that the updated ECA should create rapport with users which should result in better learning results.

8. Conclusions

The study evaluated the usability of an ECA in a CVE using heuristic evaluation. The ECA has a dual role as a tour guide and an intervention mechanism when students do not engage with the learning experience. The ECA can understand the students' affective state and call for attention when they do not engage in the learning activity. We created the behaviours of the ECA based on a theoretical framework of affect. We found that the current design of the ECA had a positive impact on the way evaluators perceived its immediacy behaviours (e.g., clarity of the ECA's verbal behaviours) and interactions. However, the evaluators agreed that additional work is needed for the ECA to properly engage students in dialogic learning scenarios. Based on the usability problems the evaluators identified, we generated a list of design recommendations on how to improve the ECA. These recommendations refer to how to improve the usability of the ECA including its immediacy behaviours (verbal and non-verbal) and interactions. When implemented, the recommendations should result in an ECA that can more effectively build rapport with users online. It is known that creating rapport relates to better learning results in both real-life and online learning scenarios. We have therefore assumed that an ECA implementing our design recommendations will have a positive impact on the student's learning performance with dialogic learning scenarios. We plan to investigate this hypothesis in a study with groups of teachers and students comparing different kinds of ECAs in dialogic learning scenarios.

Author Contributions: Writing—original draft, I.D.; Writing—review & editing, D.E.

Funding: The research that led to this paper was supported in part by the European Commission under the Contract FP7-ICT-287723 REVERIE.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bogdanovych, A.; Ijaz, K.; Simoff, S. The City of Uruk: Teaching Ancient History in a Virtual World. In Proceedings of the Intelligent Virtual Agents: 12th International Conference, IVA 2012, Santa Cruz, CA, USA, 12–14 September 2012; Nakano, Y., Neff, M., Paiva, A., Walker, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 28–35.
2. Taoum, J.; Nakhal, B.; Bevacqua, E.; Querrec, R. A Design Proposition for Interactive Virtual Tutors in an Informed Environment. In Proceedings of the Intelligent Virtual Agents, Los Angeles, CA, USA, 20–23 September 2016; pp. 341–350.
3. Zielke, M.A.; Zakhidov, D.; Hardee, G.; Evans, L.; Lenox, S.; Orr, N.; Fino, D.; Mathialagan, G. Developing Virtual Patients with VR/AR for a natural user interface in medical teaching. In Proceedings of the 2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH), Perth, Australia, 2–4 April 2017; pp. 1–8.
4. Bickmore, T.; Pfeifer, L.; Schulman, D. Relational Agents Improve Engagement and Learning in Science Museum Visitors. In Proceedings of the Intelligent Virtual Agents: 10th International Conference, IVA 2011, Reykjavik, Iceland, 15–17 September 2011; Vilhjálmsson, H.H., Kopp, S., Marsella, S., Thórisson, K.R., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 55–67.
5. Fechteler, P.; Hilsmann, A.; Eisert, P.; Broeck, S.V.; Stevens, C.; Wall, J.; Sanna, M.; Mauro, D.A.; Kuijk, F.; Mekuria, R.; et al. A framework for realistic 3D tele-immersion. In Proceedings of the 6th International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications, Berlin, Germany, 6–7 June 2013; pp. 1–8.

6. Wall, J.; Izquierdo, E.; Argyriou, L.; Monaghan, D.S.; Connor, N.E.O.; Poulakos, S.; Smolic, A.; Mekuria, R. REVERIE: Natural human interaction in virtual immersive environments. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 2165–2167.
7. Ravenet, B.; Ochs, M.; Pelachaud, C. From a User-created Corpus of Virtual Agent's Non-verbal Behavior to a Computational Model of Interpersonal Attitudes. In Proceedings of the Intelligent Virtual Agents: 13th International Conference, IVA 2013, Edinburgh, UK, 29–31 August 2013; Aylett, R., Krenn, B., Pelachaud, C., Shimodaira, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 263–274.
8. Gurtman, M.B. Exploring Personality with the Interpersonal Circumplex. *Soc. Personal. Psychol. Compass* **2009**, *3*, 601–619. [CrossRef]
9. Julmi, C. A Theory of Affective Communication. 2016. Available online: https://goo.gl/QG2K72 (accessed on 31 December 2018). [CrossRef]
10. Cone, J.G. Facilitating Trust: What Team Leaders Need to Know. Available online: http://interactionassociates.com/sites/default/files/whitepapers/Facilitating_Trust_Cone.pdf (accessed on 31 December 2018).
11. Usability Heuristics for Virtual Worlds. Available online: https://www.researchgate.net/publication/215642988_Usability_Heuristics_for_Virtual_Worlds (accessed on 31 December 2018).
12. Worley, D.; Titsworth, S.; Worley, D.W.; Cornett-DeVito, M. Instructional Communication Competence: Lessons Learned from Award-Winning Teachers. *Commun. Stud.* **2007**, *58*, 207–222. [CrossRef]
13. Gunter, G.A. The Effects of the Impact of Instructional Immediacy on Cognition and Learning in Online Classes. Available online: https://pdfs.semanticscholar.org/6ab9/88d13f9c5b747c0ee767e68b7d3eb0a05ea9.pdf (accessed on 31 December 2018).
14. Bartlett Ellis, R.J.; Carmon, A.F.; Pike, C. A review of immediacy and implications for provider-patient relationships to support medication management. *Patient Prefer. Adherence* **2016**, *10*, 9–18. [CrossRef] [PubMed]
15. Chartrand, T.L.; Bargh, J.A. The Chameleon Effect: The Perception-Behavior Link and Social Interaction. *J. Pers. Soc. Psychol.* **1999**, *76*, 893–910. [CrossRef] [PubMed]
16. Lin, L. Exploring Collaborative Learning: Theoretical and Conceptual Perspectives. In *Investigating Chinese HE EFL Classrooms: Using Collaborative Learning to Enhance Learning*; Lin, L., Ed.; Springer: Berlin/Heidelberg, Germany, 2015; pp. 11–28.
17. Hajhosseiny, M. The Effect of Dialogic Teaching on Students' Critical Thinking Disposition. *Procedia Soc. Behav. Sci.* **2012**, *69*, 1358–1368. [CrossRef]
18. Frijters, S.; ten Dam, G.; Rijlaarsdam, G. Effects of dialogic learning on value-loaded critical thinking. *Learn. Instr.* **2008**, *18*, 66–82. [CrossRef]
19. Yap, L. Module 4a Facilitation Skills Basic Instructional Skills. Available online: https://www.slideshare.net/eastleaf/module-3a-presentation-skills-basic-instructional-skills (accessed on 31 December 2018).
20. Mitchell, J.K.; Association for Talent Development. *Train the Trainer Guide*; ASTD: Alexandria, VA, USA, 2008.
21. Kuijk, F.; Apostolakis, K.C.; Daras, P.; Ravenet, B.; Wei, H.; Monaghan, D.S. Autonomous agents and avatars in REVERIE's virtual environment. In Proceedings of the 20th International Conference on 3D Web Technology, Heraklion, Crete, Greece, 18–21 June 2015; pp. 279–287.
22. Paul Lamere, P.K.; Gouvea, E.B.; Raj, B.; Singh, R.; Walker, W.; Wolf, P. The CMU Sphinx-4 Speech Recognition System. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2003), Hong Kong, China, 6–10 April 2003; pp. 2–5.
23. Muñoz, R.; Barcelos, T.; Chalegre, V. Defining and Validating Virtual Worlds Usability Heuristics. In Proceedings of the 2011 30th International Conference of the Chilean Computer Science Society, Curico, Chile, 9–11 November 2011; pp. 171–178.
24. Nielsen, J. How to Conduct a Heuristic Evaluation. Available online: https://pdfs.semanticscholar.org/6725/9adaf6e77da5082395d6d5d043fb674c08ed.pdf (accessed on 31 December 2018).
25. Nielsen, J. Enhancing the explanatory power of usability heuristics. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 24–28 April 1994; pp. 152–158.
26. Ayres, P.; Cierniak, G. Split-Attention Effect. In *Encyclopedia of the Sciences of Learning*; Seel, N.M., Ed.; Springer: Boston, MA, USA, 2012; pp. 3172–3175.

27. Calefato, F.; Lanubile, F. A Planning Poker Tool for Supporting Collaborative Estimation in Distributed Agile Development. Available online: https://www.researchgate.net/publication/235916815_A_Planning_Poker_Tool_for_Supporting_Collaborative_Estimation_in_Distributed_Agile_Development (accessed on 31 December 2018).
28. Hurst, S. Seven Ways to Increase Student Engagement in the Classroom. Available online: https://www.readinghorizons.com/blog/seven-ways-to-increase-student-engagement-in-the-classroom (accessed on 31 December 2018).

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).